R E P O R T   R E S U M E S

ED 012 889   AL 000 151

COMMUNICATION OF PRONOMINAL REFERENTS IN AMBIGUOUS ENGLISH SENTENCES FOR CHILDREN AND ADULTS.

BY- CHAI, DAVID T.
MICHIGAN UNIV., ANN ARBOR,CTR.FOR HUMAN GROWTH

REPORT NUMBER 13   PUB DATE 5 JAN 67
EDRS PRICE MF-$0.75 HC-$6.60   165P.

DESCRIPTORS- *LANGUAGE DEVELOPMENT, *PSYCHOLINGUISTICS, ELEMENTARY SCHOOL STUDENTS, JUNIOR HIGH SCHOOL STUDENTS, COLLEGE STUDENTS, RECOGNITION, LOGICAL THINKING, COGNITIVE DEVELOPMENT,

THIS RESEARCH STUDY WAS BASED ON THE ASSUMPTION THAT THE TECHNIQUE OF PSYCHOLINGUISTIC EXPERIMENTATION, AS WELL AS FORMAL ANALYSIS, CAN BE APPLIED TO DISCOVER SOME VARIABLES WHICH ALLOW THE HUMAN TO RESOLVE AMBIGUOUS SENTENCES. THE PROBLEM OF AMBIGUITY WAS CONFINED TO SITUATIONS IN WHICH A KEY PRONOUN IN A SENTENCE HAS MORE THAN ONE POSSIBLE ANTECEDENT. THE OBJECTIVE WAS TO DISCOVER SOME PARAMETERS THAT CONTROL THE ABILITY TO RESOLVE AMBIGUOUS PRONOMINAL REFERENCE IN DIFFERENT AGE GROUPS. THE EXPERIMENT INVOLVED A SET OF 176 SENTENCES, EACH OF WHICH CONTAINED PRONOUNS WITH AMBIGUOUS ANTECEDENTS. THE EXPERIMENTAL SUBJECTS--FIFTH, SEVENTH, AND EIGHTH GRADE CHILDREN, AND COLLEGE SOPHOMORES--JUDGED THE MOST APPROPRIATE REFERENT IN EACH SENTENCE. RESULTS FROM THE SOPHOMORE GROUP CONCLUSIVELY DEMONSTRATED THAT SUBJECTS COULD RESOLVE AN AMBIGUITY IN A SENTENCE BY REDUCING THE MEANING OF A KEY WORD FROM TWO TO ONE. EIGHTH GRADE SUBJECTS DEMONSTRATED THEY COULD LARGELY RESOLVE THE PRONOMINAL REFERENT WHEN REAL VERBS WERE USED (AS OPPOSED TO NONSENSE VERBS). RESULTS FROM THE SEVENTH GRADE SUBJECTS SHOWED A FURTHER DEGENERATION WITH RESPECT TO THE RESPONSES OF THE EIGHTH GRADERS. IN GENERAL THE FIFTH GRADE SUBJECTS WERE NOT ABLE TO RESOLVE THE PRONOMINAL AMBIGUITIES. THIS REPORT INCLUDES AN EXTENSIVE BIBLIOGRAPHY, THE TEST SENTENCES, RESPONSE MATRICES, AND AN "IDEAL GROUP" CHART. (AM)

THE UNIVERSITY OF MICHIGAN

CENTER FOR HUMAN GROWTH AND DEVELOPMENT
LANGUAGE DEVELOPMENT PROGRAM

Communication of Pronominal Referents in Ambiguous English Sentences for Children and Adults

DAVID T. CHAI

Report Number 13
Development of Language Functions
A Research Program-Project

Supported by:

The National Institute for Child Health
and Human Development
Grant Number 5 PO1 HD 01368-03

Administered through: *January 5, 1967*

OFFICE OF RESEARCH ADMINISTRATION · ANN ARBOR

The University of Michigan

Center for Human Growth and Development

Communication of Pronominal Referents in Ambiguous English Sentences for Children and Adults

by

David T. Chai

Report Number 13

Development of Language Functions

A Research Program-Project

January 5, 1967

Supported by the National Institute for Child

Health and Human Development

Grant Number 5 P01 HD 01368-03

ACKNOWLEDGEMENTS

The author expresses his gratitude to Dr. David McNeill for his guidance, encouragement and interest from the beginning to the end of this research; to Drs. Gordon Peterson, Arthur Melton and Richard Young and Mr. Charles Perfetti for their unique assistance; to Dr. William Livant for suggesting this topic of study during an informal discussion; to Drs. Klaus Riegel and David Birch, Co-directors of the Language Development Program of the Center for Human Growth and Development, for their encouragement and financial support; to the student Ss from the University of Michigan and the children, teachers and principals from the Adams Elementary School and West Junior High School for making the study possible; to Mrs. Judy Kingsley and Mrs. Janet Bachman for their superb and timely typing; and to his wife Audrey for her patience, encouragement and unfailing confidence without which this report could not have been completed.

The report herein was a dissertation submitted to the Horace Rackham School of Graduate Studies at the University of Michigan in partial fulfillment for the degree of Doctor of Philosophy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

In this chapter the problem of understanding the human processing of language is first discussed. As a small effort toward developing such an understanding through psycholinguistic experimentation, the present study is next outlined. Finally three related psycholinguistic studies which helped the writer in the design and analysis of the present investigation are summarized.

1.1 Defining the Problem

One primary use of natural language is to communicate ideas. To do so the ideas must be expressed in a well-structured way, according to a set of generally accepted rules, namely the grammar of a language. The set of rules is learned by children in a language community where people speak to each other. But the precise specification of these rules is very difficult. The grammar, once in the mind, allows a speaker to encode his ideas so as to be understood by others in the same community. Likewise, the grammar may be used to detect, and sometimes also correct, any deviant usage. The deviant usages may be of many different types, among which are the use of sentences with more than one interpretation.

Everyday speech abounds with ambiguous sentences. Even written material, when parsed by a computer program (usually based on a context free grammar), has been shown to contain many ambiguous constructions [Kuno and

Oettinger, 1963; Robinson, 1962]. To native speakers, however, these ambiguous constructions usually convey only one meaning. It is not known whether a native speaker without any linguistic sophistication would ever have noticed the ambiguity of such sentences. Even less known is whether a native speaker would analyze a sentence syntactically and, if ambiguous, would then apply his "semantic knowledge" to resolve the ambiguity. What is known is the fact that native speakers do often resolve syntactically ambiguous sentences, even when they occur as isolated sentences. However, for the theoretical analysis of sentences, linguists often make the simplifying assumption that a syntactic analysis precedes a semantic analysis [Chomsky, 1957, 1965; Katz and Fodor, 1963; Lamb, 1965a, 1965b]. For purposes of language processing by a computer, it is almost a necessity to make such an ordering assumption, since a computer executes its instructions serially (at least with the present generation of computers) [Simmons, 1965; Garvin, 1963].

The writer believes that the technique of psycholinguistic experimentation, as well as formal analysis, can be applied to discover some variables which allow the human to resolve ambiguous sentences. Of course the content of any manageable experiment must necessarily be narrow. (Unfortunately when the content is narrowed, the study is often criticized as trivial.) In this research study, the problem of ambiguity is confined to situations in which a key pronoun in a sentence has more than one possible antecedent. The study is part of an overall plan in the Language Development Program of the Center for Human Growth and Development to investigate the development of language behavior from children to adults. The objective in this study is to discover some parameters that control the ability to resolve ambiguous pronominal reference in different age groups. It

is hoped that once we have discovered the parameters that allow the resolution of an ambiguous pronominal reference to a certain degree, it will be possible to investigate the more difficult question of <u>how</u> the contextual information is used to resolve the ambiguity.

The experiment involves a set of 176 sentences, each of which contains pronouns with ambiguous antecedents. These sentences were given to experimental subjects (henceforth to be abbreviated by <u>S</u>, or <u>S</u>s if plural) of four age groups (fifth, seventh, and eighth grade children and college sophomores) who were asked to judge the most appropriate referent in each sentence. The controlling variables for this experiment consist of three sentence types, four grammatical forms, and four verb-pair classes. The reader is referred to Chapter 2 (EXPERIMENTAL DESIGN) which lists the variables, discusses the rationale for choosing these variables and the procedure for generating the 176 sentences from the three variables.

The resulting data will be analyzed by three mathematical models. Each model explains certain aspects of the observed behavior. The Bernoulli trials model partitions the entire set of sentences into two mutually exclusive sets, the <u>unstarred set</u> which is interpreted as containing ambiguous sentences and the <u>starred set</u> which is interpreted as containing unambiguous sentences. This model suggests a possible definition of ambiguity. The next model, the k-limited transducer model, characterizes the human as a sentence processor with limited memory in which he first decodes the sentence into its deep structure and then makes an assignment for the pronominal referent. This model explains the degenerative behavior, i.e., tendency toward random responding, when sentences become more complex syntactically or when <u>S</u>s are younger, other variables remaining constant. Finally, the information transmission model takes the response data directly

to compute various components of transmitted information (defined in Chapter 4.3). These components measure the ability to predict the pronominal referent by knowing certain combinations of variables. This model purports to explain why some combinations of variables cause a uniform response to one interpretation while other combinations cause a random response.

The proposed models are satisfactory only to the extent that they have functional equivalents, i.e., that the models account for the observed behavior. Although we can never hope to know what is happening within the central nervous system in the process of resolving pronominal ambiguity, the models will allow us to predict other unobserved behaviors which may be tested in future studies. The ultimate objective for such an investigation is to discover some of the heuristic procedures the human uses in processing and understanding sentences in natural language. If the computer is ever to become "sophisticated" enough in language processing to abstract scientific papers, retrieve key information and possibly even translate across languages, such human abilities must be encoded in the machine.

1.2 Three Related Psycholinguistic Studies

The first experiment to be reviewed is that of Kaplan [1949]. His study was an attempt to demonstrate that contextual information could be used to reduce the number of interpretations of a key word. Miller and Isard's [1964] study investigated the human capacity to recall sentences with self-embedded clauses and considered the relation of self-embedding to the theoretical model of push-down automata. The study by Fraser, Bellugi and Brown [1963] investigated the imitation, comprehension and

production of some ten grammatical features in three-year-old children. To this last study, McNeill [1965] postulated three separate memory spans to account for both the results of Fraser, et al, and of Ervin [1964] who thought that her results were contradictory to those of Fraser, et al. These studies served as guideposts from the design to the analysis of the present study.

a. Kaplan [1949]. A group of seven "translators" (some professional and some high school graduates) were presented with single words along with a set of ten meanings to be judged as appropriate, i.e., the sense in which they may be conventionally used. Then each translator was presented with the word in one of the following seven kinds of contexts: one word preceding (P1), one word following (F1), both one word preceding and one word following (B1), two words preceding (P2), two words following (F2), both two words preceding and two words following (B2) and the entire sentence. The task of each translator was to check those senses for which the test words were appropriate in each given context.

Kaplan selected 140 sentences randomly from some ten pure and applied mathematics books. The 140 key words for which the ambiguity was to be determined were the so-called "content words", i.e., nouns, verbs and adjectives. They were selected from the 140 sentences which varied in length from 15 to 40 words. For each key word, ten senses were selected from the fifth edition of Webster's Collegiate Dictionary. If less than ten senses were given by the dictionary, arbitrary senses were contrived so that every key word was given along with ten senses.

Each key word was scored according to the total number of ascriptions of acceptable senses plus the total number of denials of contrived senses.

Kaplan defined the percent reduction for a given context as the ratio of the score with that context to the score without context. However, the writer feels that the term "reduction" is better used to define the ratio of the number of senses that was eliminated in a given context to the number without context. Hence in the following paragraph, the term reduction will be used according to this new definition which differs from Kaplan's. If Kaplan gave a value of n% reduction to a word, then the new definition would give a value of (100 - n)%.

Kaplan found that P1 was the least effective in reducing the number of senses (25%), while the sentence context reduced the largest number of the senses (63%). However, the contexts P2, F2 and B2 were not significantly different from that of the sentence, which suggested that any two-words provided a context as good as the entire sentence. This classical experiment demonstrated that the reduction of senses for any key word was possible through its immediate neighboring words. It did not, however, indicate which parameters the human translators used to reduce the senses.

In the present experiment, the writer also demonstrates that reduction from two to one sense of the key word is often possible. Since he could manipulate the variables in the experiment one at a time, it was possible to discover those combinations of variables which made the reduction possible. These variables will be proposed as among the parameters we use to resolve ambiguity.

b. Miller and Isard [1964]. Twenty-four adult Ss were tested for free recall of six sentences, each 22 words long but differing in the degree of self-embedding in their phrase structures. These sentences were recorded on tape and played over earphones to one S at a time. As soon as

a sentence was heard, the S attempted to repeat it verbatim; this was recorded for analysis later. Each sentence was presented and repeated five times in this way. This continued until the S had memorized all six sentences. Six different orders of presentation of the sentences were prepared, in counterbalanced design, and four Ss learned the sentences in each order. The responses of the Ss were scored in terms of the number of words recalled in the same order as the words in the original sentence.

The different degrees of self-embedding are illustrated by the following sentences:

(0) She liked the man that visited the jeweler that made the ring that won the prize that was given at the fair.

(1) The man that she liked visited the jeweler that made the ring that won the prize that was given at the fair.

(2) The jeweler that the man that she liked visited made the ring that won the prize that was given at the fair.

(3) The ring that the jeweler that the man that she liked visited made won the prize that was given at the fair.

(4) The prize that the ring that the jeweler that the man that she liked visited made won was given at the fair.

(random) Won given liked that that the fair man made visited prize the at the the she that jeweler was the ring that.

The results of the 24 Ss are shown in Figure 1-1. The performance curves for zero and one degree of self-embedding are quite close, as are the curves for three and four embeddings, with the performance on the doubly embedded sentences falling somewhere in between. Miller and Isard explained the results through analogy with computer processing of subroutines. That is, whenever a subroutine is called during the main routine, the main routine is temporarily terminated in order to execute the subroutine. But the point of return to the main routine must be stored so that the main routine

can continue after the termination of the subroutine. But if within the subroutine, the same subroutine is called, a special kind of address storage is needed to store the return address. If this capacity for recursive calling of the same subroutine is provided in a computer program to a

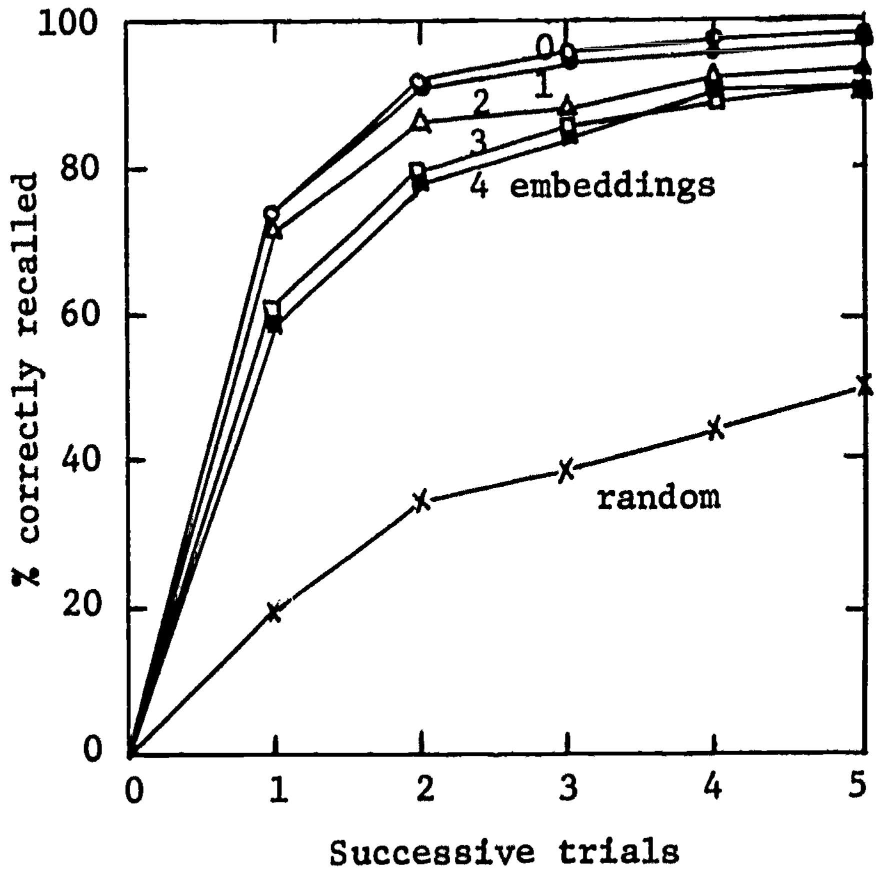

Figure 1-1: Relation of percent correctly recalled vs. successive trials.

depth of n, then the computer will be able to process (equivalent to correct recall in the human learning experiment) self-embedding sentences to a depth of n + 1. The results strongly suggest that this recursive calling capability is limited in human Ss; in fact n is one (i.e., two embeddings) if we require a 90% correctly recalled criterion after five trials.

Miller and Isard's experiment conclusively demonstrated the distinction between grammatically acceptable sentences and psychologically accept-

able sentences. Hence for practical considerations, such as writing a computer sentence parser, it is not necessary to consider sentences of indefinite degree of embedding; some small number is satisfactory, such as, say, ten.

In the present study, the k-limited transducer model will be defined and interpretations supplied to explain the degenerative behavior (i.e., tendency toward random responding) as sentences become more complex syntactically or for younger Ss. Some studies [Slobin, 1966; McNeill, 1965; Brown and Fraser, 1964] have shown that children as young as four years old have acquired the competence of both the passive and negative transformations. Hence from linguistic competence considerations, there should not be any degeneration due to either or both the passive and negative transformations in sentences, either within one age group or across the four age groups. However, it is apparent that the performance of the Ss, especially the younger children, were far inferior to their expected competence. Memory limitations are assumed to be the major factors for the inferior performance, as explained in more detail in the following section.

c. Fraser, Bellugi and Brown [1963]. A group of three-year-old Ss were tested on three tasks: imitation, comprehension and performance. Pairs of sentences were designed so that they differed only on one grammatical feature. Associated with each pair of sentences, there was a pair of contrasting pictures which distinguished the particular feature. For example, one pair of sentences, "The sheep is jumping" vs. "The sheep are jumping" was accompanied by a pair of contrasting pictures, one of which depicted a single sheep jumping over a fence while a second sheep looked on and the other depicted two sheep jumping over the fence. This pair of sentences differed only in whether the auxiliary be was singular or plural and the two pictures differed only in whether there was one or

two sheep performing the jumping. The sentence pairs with their associated contrasting pictures were used to control the performance of a child's grammar. Altogether ten different grammatical features were tested, for example, singular-plural marked by inflection ("the boy draws" vs. "the boys draw"), present progressive-past tense ("the paint is spilling" vs. "the paint spilled"), subject-object in the active ("the boy hit the girl" vs. "the girl hit the boy"), subject-object in the passive ("the car is bumped by the train" vs. "the train is bumped by the car"), etc.

In the comprehension task, the experimenter (E) read a pair of sentences to the child while pointing to the respective pictures. Then E repeated one of the two sentences and asked the child to point to the appropriate picture. He was scored right or wrong according to the picture to which he pointed. In the production task, the E again read the pair of sentences while pointing to each picture respectively. Then the E pointed to one of the pictures and asked the child to say the appropriate sentence. He was scored right or wrong only according to whether his response included the particular grammatical feature. In the imitation task, no pictures were used. The child was asked to repeat the E's model sentence and again was scored right or wrong according to whether his response included the particular grammatical feature.

The results showed that comprehension exceeded production, often by a large margin, in every grammatical contrast, while imitation exceeded comprehension on every grammatical contrast except one, again often by a large margin. Fraser, et al explained the result of comprehension exceeding production by the fact that production places a greater load on a child's memory. The imitation task depended only on perceptual-motor skills which do not operate through the meaning system, and hence it was the easiest of

the three tasks. McNeill [1965] preferred to explain the results by postulating three separate memory spans, the longest one for phonological production (equivalent to the task of imitation), the next shorter one for grammatical comprehension, and the shortest one for grammatical production. Whenever a sentence is shorter than the particular memory span, the corresponding performance is achieved correctly, but when a sentence exceeds a particular memory span, the corresponding performance cannot be achieved correctly. In the experiment by Fraser et al., the sentences were all short (with an average length of four morphemes and a maximum of eight morphemes); hence it was plausible to assume that the shorter sentences were within the grammatical comprehension span while the longer ones were within the phonological span. The only advantage of this explanation by McNeill is that it is possible to account also for the findings of Ervin [1964] who found that imitation and production were not different in her two-year-old Ss. The reason given is that Ervin's children were imitating adult speech, which was generally quite long, and that two-year-old children could easily be assumed to have shorter memory spans. Hence for these younger children, if all the sentences exceeded even the phonological span, then neither imitation nor production could be achieved accurately. Therefore, Ervin's findings were compatible with the findings of Fraser et al. through a model with three different memory spans.

Of course the specification of the size and unit of the respective postulated memory spans is a separate and very difficult question. Since it is possible for a human to encode his message to overcome his memory limitation, it is almost impossible to know the size of the span without knowing the encoding scheme. Yet the encoding scheme may differ depending on the particular task and message. If McNeill's postulated memory spans are tenable, then there are probably other spans for other tasks.

In this experiment, two memory spans are postulated to account for the observed degenerative behavior as sentences become more complex syntactically or as Ss become younger. The model used is a concatenation of two k-limited transducers, one for decoding the input sentence into its deep structure and the other for taking the deep structure as input to assign the pronominal referent. The k for the k-limited transducer reflects the maximum number of input symbols that the transducer can process, and upon interpretation, the k reflects the memory span of the S. When a sentence falls within a S's decoding span and in addition the decoded message is within the S's assignment span, then the referent will be assigned correctly. However, as a sentence becomes more complex syntactically, the chance of exceeding a S's decoding span becomes greater. Likewise as a S becomes younger, it is reasonable to assume that both the decoding and assignment spans become shorter, and hence the chance for a sentence to have an incorrectly assigned referent becomes greater.

Chapter 2

EXPERIMENTAL DESIGN

This chapter describes the construction of the test sentences. The controlling variables and the rationale for these variables will be presented first; then the procedure for generating the test sentences will be explained.

2.1 Controlling Variables

The experiment consisted of a set of 176 sentences, all of which could have two interpretations. The sentences were constructed from the following three variables: three sentence types, four grammatical forms, and four sets of verb-pairs.

The three sentence types are:

John Yed Bill and/but he Xed him.
John Yed Bill and/but he Xed him back.
John and Bill Yed each other and/but he Xed him.

The four grammatical forms are:

Active (A), e.g., he Xed him.
Negative ($\overline{A}$), e.g., he did not X him.
Passive (P), e.g., he was Xed by him.
Negative-passive ($\overline{P}$), e.g., he was not Xed by him.

The four classes of verb-pairs are:

Nonsense verbs (NS), i.e., Y and X are consonant-vowel-consonant syllables.
Identical verbs (ID), i.e., Y and X are the same verb, from the set: {tease, hurt, strike, punch}.

Similar verbs (SI), i.e., Y and X are both outward action verbs, from the set of ordered pairs: {<hit, kick>, <push, shove>, <help, assist>, <frighten, scare>}.

Logical reversal verbs (LR), i.e., Y is an inward action verb, and X is an outward action verb, from the set of ordered pairs: {<understand, answer>, <hear, call>, <remember, phone>, <recognize, invite>}.

For example, the following three sentences were constructed by using the nonsense verb-pairs:

Bill qeged John and he yeced him. [A-A]
John was zeved by Bill and he pehed him back. [P-A]
Bill and John did not kez each other but he was qemed by him. [$\bar{A}$-P]

The first sentence was constructed by using the first sentence type, with the first clause in the active grammatical form (A) and the second clause also in the active grammatical form (A). The second sentence was constructed by using the second sentence type, with the first clause in the passive grammatical form (P) and the second clause in the active grammatical form (A). The third sentence was constructed by using the third sentence type, with the first clause in the negative grammatical form ($\bar{A}$) and the second clause in the passive grammatical form (P).

Further, the following three sentences were constructed by using the real verb-pairs for each of the three sentence frames:

John did not hurt Bill but he hurt him. [$\bar{A}$-A]
Bill did not help John but he was assisted back by him. [$\bar{A}$-P]
John and Bill were not understood by each other and he did not answer him. [$\bar{P}$-$\bar{A}$]

The first sentence used one member from the class of identical verb-pairs in its $\bar{A}$-A construction; the second sentence used one member from the class of similar verb-pairs in its $\bar{A}$-P construction; and the third sentence

used one member from the class of logical reversal verb-pairs in its $\bar{P}$-$\bar{A}$ construction.

2.2 *Terminology*

The term *sentence type* will be applied to the three sentence structures where X and Y can be any verbs, and the two clauses can be in any one of four grammatical forms. The three sentence types will be abbreviated respectively as the *He Xed him* type, *He Xed him back* type and *Each other* type. A *sentence frame* will be any one of the sentence types with a particular set of verb-pairs. Hence there are four sentence frames for each sentence type. When a specific sentence frame is referred to, the particular sentence type will be preceded by an adjective which describes one of the four verb-pair classes. For example, the nonsense *He Xed him* frame is that frame in which the verbs X and Y are nonsense syllables in the *He Xed him* sentence type. Since every sentence is constructed out of two clauses, each can be in one of four grammatical forms, a specific sentence of a sentence type or sentence frame will be referred to by its grammatical forms. For example, the A-P sentence of the *He Xed him* type is *John Yed Bill and he was Xed by him*, and the $\bar{A}$-P sentence of the identical *He Xed him back* frame is *John did not strike Bill but he was struck back by him*.

2.3 *Rationale for the Variables*

Of the three sentence types, the *Each other* type was intended to serve as a basis of comparison. It was expected that any sentence of this type would be so ambiguous that it would produce random responding and thus provide a base line against which to compare responses to the

other two sentence types. The *He Xed him back* type was used because it contains probably the most frequently exposed compound sentences for children. The referent for the pronoun *he* should be unambiguous because the word "back" provides a semantic context which leads the reader to switch the subjects of the two verbs in the two respective clauses. Thus if *John Yed Bill*, then it must be *Bill* who *Xed him back*. In the *He Xed him* type, the most plausible interpretation is also to switch the subjects of the two verbs in the two respective clauses. Here the subtle contextual information comes from the structure of the two clauses; that is, if *John Yed Bill* and *John Xed him* are implied, then there exists another construction which would better express the two relations, as for example, the construction *John Yed and Xed Bill*. However, if the verbs X and Y were the same, then the construction *John Xed and Xed Bill* (e.g., in the context, *John kicked and kicked and kicked . . . Bill*) has certain poetic flavor for emphasizing the action, but is rarely used in ordinary writing or conversation. Hence in this situation, the more plausible interpretation would be *John Xed Bill and Bill Xed him*.

For each of the two clauses in a sentence type, four grammatical forms were used: active (A), negative ($\overline{A}$), passive (P), and negative passive ($\overline{P}$). Hence each sentence frame had a set of 16 individual sentences. These sentences were different ways of expressing the same idea, but were expected to vary in the degree of ambiguity of the pronominal referent (as measured by the non-uniformity of response from groups of *S*s). It was not known initially whether the use of the conjunction *and* or *but* would alter the degree of ambiguity, but the following rule was consistently used: *and* was the conjunction whenever both clauses were affirmative (i.e., A or P), or both negative (i.e., $\overline{A}$ or $\overline{P}$), and *but* was used whenever one clause was affirmative and the other was negative.

Since every sentence was made up of two clauses, it seemed intuitive that altering the relationship between the two verbs could alter the interpretation of the sentence, and hence the choice of the pronominal referent. Since there were two persons and two verbs in each sentence, there were four possible ways of expressing the two actions in the sentence types He Xed him and He Xed him back:

1) John Yed Bill and John Xed Bill (back).
2) John Yed Bill and Bill Xed John (back).
3) Bill Yed John and Bill Xed John (back).
4) Bill Yed John and John Xed Bill (back).

For any given sentence, the first clause was already specified, hence the number of interpretations was reduced to two (of course one could introduce other possibilities by equating him to someone other than John or Bill.)

Three classes of verb-pairs were taken from a previous study [Chai, 1966] in which the Ss were instructed to insert two real verbs before they decided on the interpretation of the pronoun he. Each class of verb-pairs made one interpretation more plausible than the other. The identical verb-pairs made interpretations one and three almost impossible; the logical reversal verb-pairs made interpretations one and three most likely; while the similar verb-pairs were less determinant.

The class of logical reversal verb-pairs was constructed in such a way that the first verb always implied a previous action in which the second verb was a response. The use of "reversal" denotes a change of subject-object relationship. For example, John understood Bill implies that Bill asked (questioned, called, etc.) John; consequently, in order to decide on the referent of he in the sentence John understood Bill and he answered him (back), the S had to rely on some implied sentence, as, for example, Bill questioned John and he answered him (back). Once the

implied sentence was constructed, interpretations two and four were almost certain, i.e., he = John. Hence in the original form, John understood Bill and he answered him (back), the fact that he = John corresponded to interpretations one and three. For the other verb-pairs in the logical reversal class, the verb hear implied most likely the verbs ask, yell, call, etc; the verb remember implied most likely the verbs or verb phrases call, leave a message, make an appointment, etc; the verb recognize implied most likely the verbs meet, call, surprise, etc.

The class of identical verb-pairs was constructed from verbs that could be repeated in a sentence, e.g., John hurt Bill and he hurt him (back). The class of similar verb-pairs was constructed from sets of two verbs which denoted very closely related actions, e.g., <hit, kick>. The order of the two verbs in the class of similar verb-pairs is assumed to have no difference in the choice of the pronominal referent, e.g., John hit Bill and he kicked him (back) and John kicked Bill and he hit him (back) have the same referent for he. But such is not the case for the class of logical reversal verb-pairs, e.g., John understood Bill and he answered him (back) and John answered Bill and he understood him (back) have very different referents for the pronoun he. All the verb-pairs in the above three classes, except <remember, phone> and <recognize, invite>, were selected from the set of verb-pairs that a group of Ss supplied in the previous study.

Nonsense syllables formed another class of verb-pairs. The purpose of using the nonsense words was to investigate the transmission of pronominal reference by the structure of the sentence rather than by means of extra-linguistic information. Presumably the sentence with nonsense verb-pairs would be responded to least uniformly, especially by the younger children. These nonsense verb-pairs were randomly selected from a set of

consonant-vowel-consonant syllables having a Krueger index from 50 to 59 [Krueger, 1934]. These belonged to the set with a relatively small number of associations. The Ss were specifically instructed not to substitute real verbs for the nonsense ones, but to decide the pronominal referent solely on the structure of the sentence. Of course there was no guarantee that the Ss did not supply their own lexical meaning for the nonsense verbs before they decided on the pronominal referent. But the fact that sentences with nonsense verbs were responded to less uniformly than the same sentences with real verbs indicated that the Ss had more difficulty with this set of sentences. In addition, the average time for responding to the nonsense sentences (11.3 seconds for college sophomores) was longer than sentences with real verbs (10.5 seconds), but much less than the average time for writing a pair of real verbs before responding to sentences with nonsense verbs, as estimated from the previous study (24 seconds).

2.4 Sentence Generation Procedures

For each sentence frame, 16 sentences were constructed from all combinations of active (A), negative ($\overline{A}$), passive (P) and negative-passive ($\overline{P}$) grammatical forms of the two individual clauses. Ideally each verb-pair class needed 16 members, but it was almost impossible to construct that many verb-pairs within the vocabulary range of the fifth graders (except in the case of the nonsense verb-pair class) without repetition and without altering the class interpretation. Thus, the decision was made to use four words for each of the real verb-pair classes. A "Latin square" procedure as shown in Table 2-1 was used in assigning the verb-pairs to the individual sentences of a sentence frame. The rows of the matrix in Table 2-1 correspond to the grammatical form (A, $\overline{A}$, $\overline{P}$ or P) of the first clause while the columns correspond to the grammatical form (A, $\overline{A}$, $\overline{P}$ or P)

of the second clause. Each cell, then, corresponds to a particular sentence. For example, the $\bar{A}$-P cell (row $\bar{A}$ and column P) corresponds to the sentence *John did not Y Bill but he was Xed by him*. The numbers (1 to 4) in each cell correspond to the first to fourth verb-pair of

Table 2-1: "Latin square" for the assignment of verb-pairs.

	A	$\bar{A}$	$\bar{P}$	P
A	1	3	2	4
$\bar{A}$	2	4	1	3
$\bar{P}$	3	1	4	2
P	4	2	3	1

each of the three classes of real verb-pairs. Thus for the *He Xed him* sentence frames, the number 3 in the above $\bar{A}$-P cell indicates that the third verb-pair from each verb class was substituted for the verbs Y and X in the sentence *John did not Y Bill but he was Xed by him*:

Identical: John did not strike Bill but he was struck by him.
Similar: John did not help Bill but he was assisted by him.
Logical Reversal: John did not remember Bill but he was phoned by him.

The following were the number of sentences thus generated. *He Xed him* and *He Xed him back* sentence types each had four classes of verb-pairs and 16 different grammatical forms. Thus there were $2 \times 4 \times 16 = 128$ sentences. In addition, the *Each other* sentence type had three classes of verb-pairs (the identical verb-pairs were not allowed because

it was felt that some of the resulting sentences would be anomalous, or even ungrammatical, as for example, John and Bill did not strike each other but he struck him) and 16 different grammatical forms, giving 48 sentences. Thus there were all together 128 + 48 = 176 sentences.

For each sentence, the choice of John or Bill as the grammatical subject was random. The set of 176 sentences were grouped into two parts, one containing 48 sentences that had nonsense syllables as the verb-pairs, the other containing 128 sentences that had the real verb-pairs. Within each part, the sentences were typed in a random order. The sentences are reproduced in Appendix A.

Chapter 3

DATA

In this chapter the data collection procedures will be described. First, the considerations involved in choosing the age groups will be discussed, then the procedural instructions for the Ss will be presented, and finally the preparation of the resulting data for computer analysis will be explained.

3.1 Choice of Age Groups

The data for the investigation consisted of responses from four age groups. The age groups were selected with the aim of covering the period of transition from near inability to resolve the pronominal referent to near uniformity in the choice of the pronominal referent. Even though some psychologists have shown that children of four to six years can command the passive and negative transformation fairly well [Slobin, 1966; Brown and Fraser, 1964; McNeill, 1965], much more linguistic competence is required to select the appropriate pronominal referent. For example, it was necessary to generalize from the everyday sentence Johnny hit me and I kicked him back to John bejed Bill and he (Bill) gowed him back and further to John was not zeved by Bill but he (Bill) was kezed back by him. It was decided to use the written form, rather than the oral form, for the following reasons: some of the sentences were fairly long and were expected to exceed the immediate memory span of younger children; some of the nonsense syllables were difficult to pronounce; it was impossible to

decide how many times a sentence should be repeated orally to insure comprehension and whether the number should be constant for all sentences and/or all age groups; and finally, individual oral testing would have been too time consuming, whereas group oral testing would have been too difficult to administer.

Once the written form was decided upon, considerations of reading ability required that the youngest subjects be at least in the fourth grade. Finally fifth grade was selected as containing the lowest age group that would reasonably well insure adequate reading ability and interest. Interest was a necessary criterion, since the testing was to occur during regular class hours. Seventh and ninth grades were also selected for the tests, and in addition college sophomores were chosen to represent the adult language population. Unfortunately, the time for testing the ninth grade was during the last class of the day, which apparently led to much random responding in an attempt to finish quickly. Subsequently, a large group of eighth graders were tested and to be tested again three months later to see if these children would respond differently (hopefully more uniformly) after they were introduced to some modern concepts of structural linguistics. The results from the first testing (before structural linguistics) for the eighth grade children were used in place of the ninth graders because their average IQ's match more closely to the seventh graders. However, the responses of the ninth graders are presented in Appendix B of this report.

The fifth grade children came from the Adams School and the seventh, eighth and ninth grade children came from the West Junior High, all of which are public schools in Ypsilanti, Michigan. The college Ss were students at the University of Michigan and were paid. The children were tested for

approximately 40 minutes in their regular class session, without any monetary compensation. The college Ss were tested in two groups, one in the afternoon and the other in the evening.

3.2 Procedural Instructions for the Subjects

The testing materials were prepared in the form of a booklet with written instructions. The experimenter (the writer) always described briefly the purpose of the experiment to the Ss and explained the instructions with some specific examples. The booklet was divided into two parts, one containing 48 sentences that had nonsense syllables as the verb-pairs and the other containing 128 sentences with real verb-pairs. Following the normal counterbalancing procedure, half of the Ss in each age group worked the set of nonsense verb-pairs before the set of real verb-pairs, while the other half worked in the reverse order. They were instructed to record the beginning and terminating time for each part.

The Ss were instructed to mark an "x" above the word John or Bill to denote the most appropriate referent of he. They were further instructed to answer every sentence in the given order. One or two Ss who responded with an "x" on the word for the first mentioned person (i.e., the first word) in every sentence on at least five of the seven pages were discarded from the analysis. In the children's groups, some of the Ss were not able to finish during the class hour; hence the set containing the unfinished sentences was thrown out. The result was that the total number of Ss was reduced for some of the frames. In particular the eighth grade had one fewer male S (25 male Ss instead of 26), the seventh grade had one fewer female S (12 female Ss instead of 13), and the ninth grade had one fewer

male S (13 male Ss instead of 14). All discarded sentences contained the set of nonsense verb-pairs. Hence for the nonsense He Xed him frame, the nonsense He Xed him back frame and the nonsense Each other frame, the total number of male eighth grade Ss was 25, the total number of female seventh grade Ss was 12 and the total number of male ninth grade Ss was 13.

3.3 Data Collection

The results were transcribed by hand to a coding sheet. A response to the grammatical object-slot (i.e., the mark "x" was above the word that was the grammatical object of the first verb) was recorded as 1 and a response to the grammatical subject-slot was recorded as 0. For the Each other sentence type the first mentioned person was taken as the subject and the second mentioned person as the object. This was only intended as a notational convenience so that the terms "subject" and "object" could be used for all three sentence types. These coding sheets were then punched and verified. A MAD program was written to accept these data cards to tabulate the number of 1's (object-slot) for each group of Ss (males and females separately) and for each of the 16 sentences in a sentence frame. These results were punched out as data for an information transmission analysis program. The results for the college sophomores, for example, appear as 4 x 4 response matrices in Tables 5-1, 5-2 and 5-3. The number in each cell corresponds to the number of Ss who responded with the object-slot. Hence the difference between the number in each cell and the total number of Ss is the number of responses with the subject-slot. The "total" matrices are pooled data from both sexes.

Chapter 4

MODELS OF ANALYSIS

In this chapter three mathematical models are proposed for analyzing the observed data. Each model will be defined and interpretations will be supplied to relate the experiment with the mathematical model. The Bernoulli trials model partitions the entire set of sentences into two subsets, one containing the unambiguous sentences, the other containing sentences with ambiguous pronouns. The k-limited transducer model characterizes the human as a sentence processor with very limited memory. This model explains the degenerative behavior (tendency toward random responding) when sentences become more complex syntactically or when Ss are younger, other variables remaining constant. Finally, the information transmission model characterizes the Ss' task of reading each sentence and giving a response for the pronominal referent as a communication system. It defines the term "transmitted information" as a measure of the degree of predictability of the pronominal referent from the structure of the sentence.

4.1 Bernoulli Trials Model

Feller defined Bernoulli trials as "repeated independent trials, each with only two possible outcomes, and their probabilities remain the same throughout the trials" [1957, p. 135]. In other words, we let there be n trials, each with probability p in obtaining the outcome 1, and probability $1 - p$ in obtaining the outcome 0. Let the random variable

X_i represent the outcome of the i^{th} independent repeated trial, $i = 1,2, . . . , n$. Then the random variable $Y = X_1 + X_2 + . . . + X_n$ represents the outcome of Bernoulli trials of n independent events (or trials). The random variable Y has a binomial distribution which is tabulated for different values of p and n (e.g., Tables of the Cumulative Binomial Probability Distribution, Harvard University Press, 1955).

We can then calculate the probability of obtaining an observed value of Y in an n-trial. But for a large value of n, this probability is always very small; hence it is better to calculate a confidence interval $[u,v]$ such that for a given choice of α, the probability of $Y \leqslant u$ or $Y \geqslant v$ is $< \alpha$. For this experiment, α is chosen to be .01. Hence the probability of obtaining a Y within the confidence interval is $1 - 2\alpha$ or .98 when we assume all the Ss in a group responded in random way, i.e., that each response is a Bernoulli trial. Hence if Y falls outside the confidence interval $[u,v]$, then it is unwise to assume that Ss' responses can be described by Bernoulli trials.

The Bernoulli trials model is used in the present experiment to separate the sentences which have ambiguous pronominal referents from those that do not. The probabilistic approach allows us to select a certain chance level α of being wrong. In the following we will describe our experiment in relation to the model, and will employ a statistical method for estimating the value of p.

The task of each S in this experiment was to select the most appropriate referent for the pronoun he in each sentence. Since he was instructed to mark an "x" on top of either John or Bill as the referent of he, the S had only two choices as he read each sentence. There is a certain probability p that John would be selected along with the probabil-

ity 1 - p that Bill would be selected. Hence the outcome of a single S responding to a particular sentence is a trial with two outcomes. Instead of assuming p to be .50, we will describe a method later to estimate the value of p from the group responses so that it will represent each S's probability of responding with John. Since it is reasonable to assume that each S in a group responds independently of any other in the same group, we have a case of Bernoulli trials.

Since by experimental design, the choice of John or Bill as the grammatical subject of each sentence was random, it would be more revealing to consider the outcome, not as John or Bill, but as a grammatical subject-slot or a grammatical object-slot. Thus we let the random variable $X_i = 1$ when the outcome is an object-slot, and $X_i = 0$ when the outcome is a subject-slot. Then the random variable $Y = X_1 + X_2 + \ldots + X_n$ is the number of Ss out of n who responded with the object-slot. The numbers in the response matrices such as those in Tables 5-1, 5-2 and 5-3 are the values of Y for each sentence in the college sophomore group. The total matrices are the cell-by-cell addition of the corresponding male and female responses. For each value of Y, there is an implied number $n - Y$ which corresponds to the number of Ss who responded with the subject-slot.

The interpretation of the confidence interval is the following. If a sentence has an ambiguous pronoun he, then each S, in a forced choice condition, has to select a referent arbitrarily. There may, nevertheless, be some bias toward the grammatical object (or subject) slot so that the value of p may not be .50. Whatever the value of p, the sum of the individual random responses for a group of Ss follows a binomial distribution. Hence if an observed Y falls within the confidence interval [u,v], we have a 98% chance that the corresponding sentence elicitated random re-

sponding due to its ambiguous pronoun. On the other hand, if a sentence has an unambiguous pronoun, then each S would have responded with the same referent. Hence the resulting value of Y is either very large, $\geqslant v$ (when the referent is the object-slot) or very small, $\leqslant u$ (when the referent is the subject-slot). Since the value of Y is outside the confidence interval $[u,v]$, we would not expect it to have resulted from random responding by the Ss. The confidence interval serves to partition the set of sentences into two subsets, one containing those sentences with an ambiguous pronoun, the other containing those sentences with an unambiguous pronoun. Therefore the Bernoulli trials serve as an operational definition for the concept of pronominal ambiguity.

The value of p is estimated in the following way. Since each age group, each sentence type, and each set of verb-pairs may cause a certain bias toward the object (or subject) slot, we are forced to calculate p separately for each sentence frame in an age group. Since all the 16 sentences in a frame are different ways of expressing the same idea, they provide 16 sample observations by which we can estimate p. Let these values be denoted as Y_i, $i = 1,2, \ldots ,16$, which corresponds to the number of Ss (from a total of n Ss) who responded with the grammatical object-slot to the 16 sentences of the frame. Since these numbers are the results of Bernoulli trials with unknown p, we can estimate this p by the maximum likelihood estimator [Mood, 1950, Chapter 8],

$$\hat{p} = \frac{\sum_{i=1}^{16} Y_i}{16\, n}$$

The numbers in each response matrix such as those shown in Table 5-1 are the number of Ss who responded with the grammatical object-slot to

the 16 sentences of a sentence frame, i.e., they are the Y_i's. The maximum likelihood estimator $\hat{p}$ and the confidence interval [u,v], where α is chosen as .01, are written immediately under each matrix. Those Y_i's which fall outside the confidence interval [u,v] are starred. The sentences corresponding to the starred cells are interpreted as unambiguous sentences.

4.2 K-limited Transducer Model

The following formal model is used to account for the degenerative behavior (tendency toward random responding) both within a sentence frame when sentences become more complex and across age groups when Ss are younger. Correct responding depends on the Ss having decoded the sentences into their deep structures and on some general knowledge of the use of the English language in describing two actions between two persons. We can use a k-limited transducer that takes as input the given sentences and

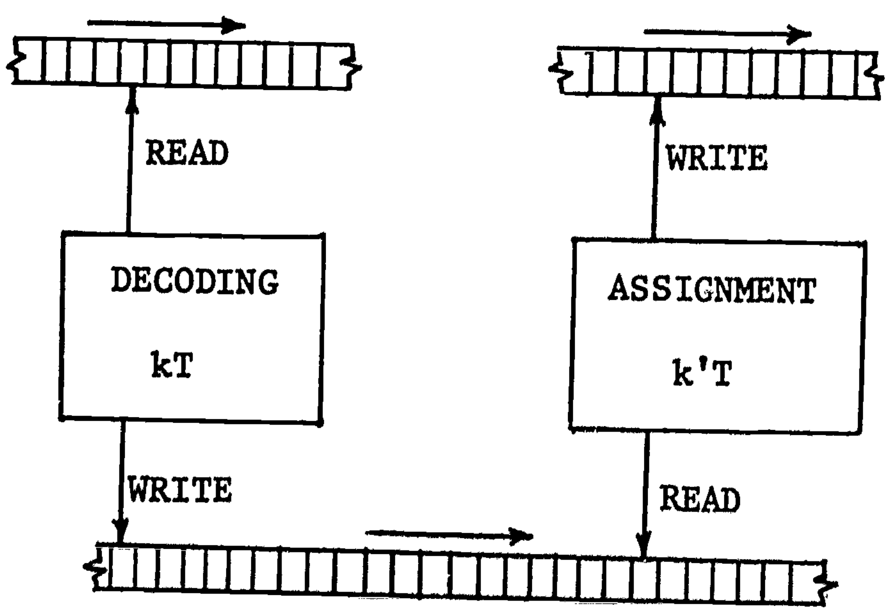

Figure 4-1: Model of human sentence processor. The kT limits the decoding of a sentence and the k'T limits the assignment of the pronominal referent.

produces as output the deep structures of the input sentences. The output is then fed to another k-limited transducer which uses the general rules of English usage to assign an output of John or Bill. A diagram of the k-limited transducer model is shown in Figure 4-1.

A k-limited transducer kT is a device that can read an input tape and write an output tape. Let A_I, A_0 and Σ be finite sets of input symbols, output symbols and internal states respectively. Then the transducer kT is completely specified by a set of transition rules of the form:

$$\langle x, \Sigma_i\rangle \rightarrow \langle y, \Sigma_j\rangle$$

where

x is a string over the input symbols A_I and its length $lg(x) < k$,

y is a string over the output symbols A_0,

Σ_i is an i^{th} internal state,

Σ_j is a j^{th} internal state.

The rule specifies that when the kT is in its state Σ_i, reading an input string x of at most k units, the kT writes the string y on the output tape while its internal state is moved to another state Σ_j. During this process, the input tape moves to the right $lg(x)$ units so that the kT is ready to read the next string of at most k units. Likewise, the output tape also moves to the right $lg(y)$ units so that the kT is ready to write the next output string. The kT is a more general transducer than the finite transducer defined by Chomsky [1963, p. 346], which is an 1-limited transducer.

The intended interpretation for the k-limited transducer in the present experiment is to consider each S as a sentence processor. The S decodes each clause of the sentence into its deep structure and uses his

knowledge of English (i.e., his linguistic competence) to assign the pronominal referent. The input and output symbols may be taken as the words. As for the decoder stage, there is a certain limitation up to which the S can process correctly. Since the task was to identify the referent of the pronoun he, the S must first discover the subject-object relation in the two clauses so that the two actions can be connected with John and Bill. When a clause is expressed in the active form, there is very little processing needed to identify the subject and the object, since the surface structure is the same as the deep structure. When a clause is in the negative form (still active), the subject-object relation has not changed from the affirmative form, and for the present task, the negation is not expected to affect the choice of the pronominal referent. However, when a clause is in the passive form, the subject-object relation in the surface structure is just the opposite of that in the deep structure. Hence a certain amount of processing is needed to obtain the relationship between the logical subject and the logical object. The k in the kT reflects not only the length of the clause but also the "program" which obtains this relationship of the sentence. Thus, although the length of a clause is the same no matter which class of verb-pairs is used, the "program" for the real verb-pairs could be shorter than for the nonsense verb-pairs. Therefore, given a certain decoding capacity k for a S, if the length of the clause and the demands of the processing "program" exceed this k, then that clause will be erroneously decoded. We do not know how the S stored the decoded message, but the subject-object relationship and the particular verb used must be clearly stated. Since each sentence consists of two clauses, the decoding process can proceed on one clause at a time. For example. the four grammatical forms in the first

clause may be decoded as follows:

John Yed Bill	$\rightarrow$	Yed [John, Bill]
John did not Y Bill	$\rightarrow$	not Yed [John, Bill]
John was Yed by Bill	$\rightarrow$	Yed [Bill, John]
John was not Yed by Bill	$\rightarrow$	not Yed [Bill, John]

The decoded message shows clearly the logical-subject and the logical-object of the clause within the square bracket. For the second clause, however, we need only know the logical function (subject or object in the deep structure) for the pronoun <u>he</u> and the type of sentence. Hence the eight different clauses may be decoded as follows:

he Xed him (back)	$\rightarrow$	Xed (back) [subject]
he did not X him (back)	$\rightarrow$	not Xed (back) [subject]
he was Xed (back) by him	$\rightarrow$	Xed (back) [object]
he was not Xed (back) by him	$\rightarrow$	not Xed (back) [object]

For example, the sentence

John Yed Bill but he was not Xed back by him

has the following as its decoded message,

Yed [John, Bill] but not Xed back [object].

The decoded message is then fed to a second transducer $k'T$ where the rules of English usage are applied to assign an output, either <u>John</u> or <u>Bill</u>. The k' reflects the maximum length of the decoded message plus a "program" which assigns the pronominal referent. Again it is necessary to postulate an assignment "program" which is concatenated with the decoded message, because the difficulty in making the assignment varies with the class of verb-pairs and the sentence type. Thus the level of difficulty is reflected in the length of the "program" for making the assignment. Hence, if two given sentences can be easily decoded but one is more difficult to assign than the other, then we can expect that the more difficult sentence will not receive a uniform response from a group of <u>Ss</u>. In terms

of the k-limited transducer model we say that both sentences are within the k of the decoding transducer, but the more difficult one has exceeded the k' of the assignment transducer due to the longer assignment "program".

There is some behavioral evidence in support of including the "program" as part of the input string of symbols. Mehler [1963] and Mehler and Miller [1964] found that when recalling English sentences, the most frequent errors were of a syntactic type, i.e., errors of recalling a grammatical form different from the one given. The authors suggested that the Ss decoded the given sentence into its kernel form and tagged it with the necessary transformational rules. Savin and Perchonock [1965] performed an experiment in which the S was presented with a sentence followed by a sequence of eight unrelated words. The task was to recall the sentence verbatim and as many of the unrelated words as possible. Then another sentence and a sequence of eight different words was presented to the S for recall. The sentences differed from each other by a number of grammatical transformations, e.g., the passive, the question, the negation, the emphasis, and the wh transformations. The hypothesis was that if the Ss were to decode each sentence into its kernel form plus tags indicating the necessary transformations, then for a given memory capacity the more complex sentences would take up more "space", and as a consequence there would be less "space" to remember the additional unrelated words. Indeed, the number of additional words recalled correctly was greatest for the kernel sentences, and decreased for those that had more transformations. In other words, the "program" for encoding the decoded kernel sentences was longer for the more syntactically complex sentences, i.e., those with more transformations.

The experimental results of Epstein [1961] and Rosenberg [1966] support the concept of including a "program" as inputs to the assignment k-limited transducer. Epstein found that semantically anomalous, yet grammatical, sentences were learned faster than the same sentences with nonsense words. Rosenberg found that for both grammatical and ungrammatical sentences, there was a significant difference in recalling the sentences when the association levels were changed for the content words. In the present experiment, we have four classes of verb-pairs for each sentence type. Thus we would expect the sentences with nonsense verb-pairs to be the most difficult, those with identical verb-pairs to be the easiest, and those with either the similar verb-pairs or logical reversal verb-pairs to fall somewhere in between. To characterize this level of difficulty, which is not simply the length of sentences, we hypothesize an "assignment program" in parallel to that of a "decoding program".

4.3 Information Transmission Model

An information transmission model may be diagrammed as in Figure 4-2. The input sources $A, B, \ldots, N$ have $\bar{\bar{A}}, \bar{\bar{B}}, \ldots, \bar{\bar{N}}$ symbols respectively, and the output source Z has $\bar{\bar{Z}}$ symbols. The symbols from the input sources are encoded into a message which is transmitted to the destination. At the destination, the received message is decoded to produce an output symbol. In the transmission process, there is a certain amount of noise introduced which makes it impossible to predict the output symbol with complete accuracy from the input symbols.

We can interpret the task of each S̲ through the information transmission model in the following way. There are three input sources, A, B and C. Source A has four symbols A, $\bar{A}$, $\bar{P}$ and P which denote the

grammatical form of the first clause of a sentence; source B also has four symbols A, $\overline{A}$, $\overline{P}$ and P which denote the grammatical form of the second clause of the sentence; source C has two symbols M and F which denote the sex of the S who is responding to the sentence. Each sentence of a sentence frame is thus an encoded message from these three

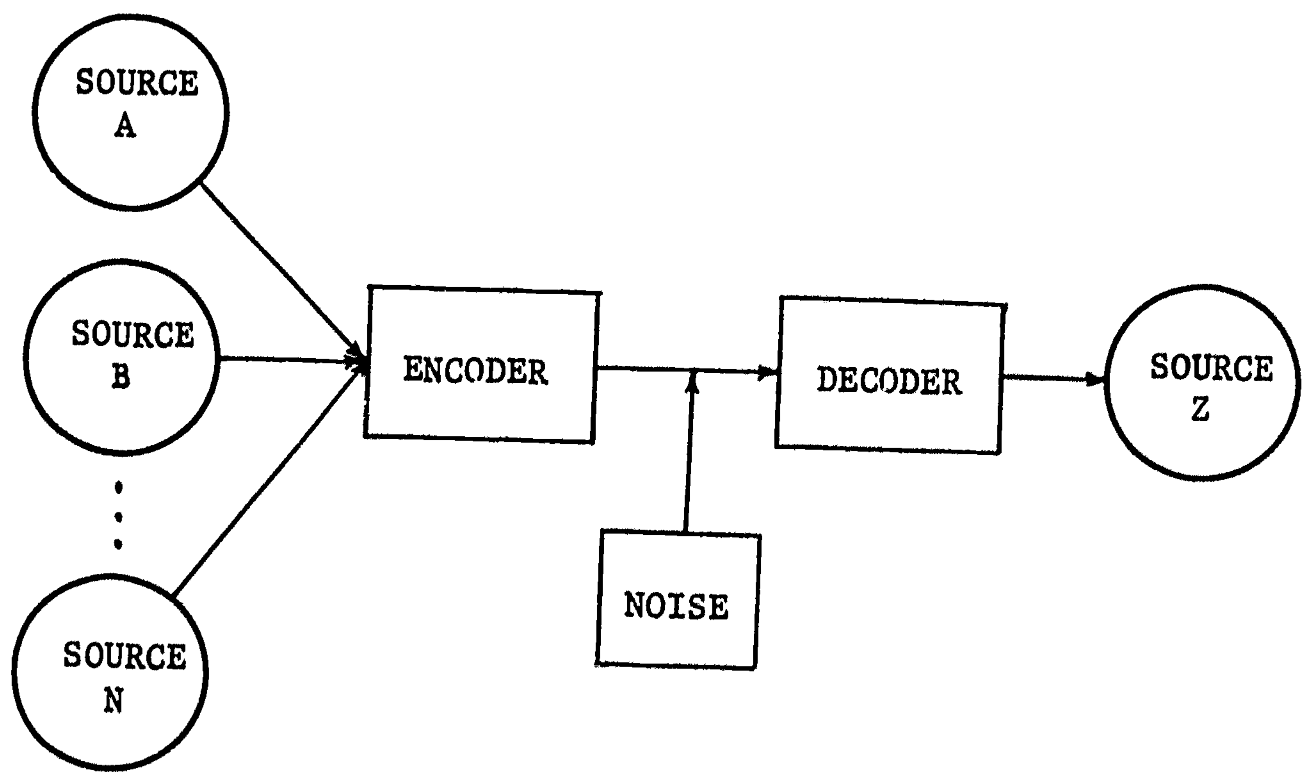

Figure 4-2: Diagram of an information transmission model.

input sources. The encoded message is read and decoded in the brain from which an output symbol is produced. The output source consists of two symbols, John and Bill or subject and object or 0 and 1. In the process of transmission from the input symbols to the output symbols, a certain amount of noise is produced so that the output symbols are not completely dependent on the input symbols. In the present experiment, many different factors could have contributed to the non-uniformity of the Ss' responses (i.e., the noise component), among which are mistakes in decoding sentences, nonequivalent linguistic knowledge among the Ss, insufficient

cues in the sentence structure (as in the sentences from the Each other type), or insufficient memory capacity to perform this task (as in the fifth grade group). Such internally generated noise is not controllable, but there may still be sufficient dependency between the input and the output sources to allow a prediction of the output symbols that is significantly better than chance.

Other variables, such as the verb-pairs, the sentence types and the age groups could have been included as input sources. But since the measure of "information", which is defined later, is the average value over a random variable, it would be meaningless to calculate such an average value when we expect very different responses among the elements of the random variable. For example, we expect that the Each other sentence type would elicit random responding while the other two sentence types would elicit uniform responding, especially for the older groups. Hence by averaging the responses from the three sentence types, we lose all the distinctions among the sentence types. Similarly, because the logical reversal verb-pairs are expected to elicit responses opposite (i.e., reversed) from those of the identical verb-pairs, that distinction would also be lost if the verb-pairs variable were averaged. Also, we expect the different age groups to respond differently so that an average response from the four age groups would not represent any population group. However, when we treat each sentence frame as a unit, the 16 sentences are different ways of expressing the same idea. Therefore, the average of these 16 responses (represented by the two variables A and B) is the measure of the degree of ambiguity of that sentence frame. The sex is taken as a variable because we want an average response for each age group, especially when the male and the female responses do not differ significantly.

Before we turn to a mathematical derivation of a 3-input source and 1-output source information transmission model, we will give an informal outline of the procedure and the important concepts that are involved. The derivation begins with a 1-input source and 1-output source model which is exemplified by a 2-dimensional table where the rows correspond to the input symbols and the columns correspond to the output symbols. Each cell in the table corresponds to the event that the input symbol as designated by the row has elicited the output symbol as designated by the column. The number in the cell denotes the frequency with which that event has occurred. With such a contingency table, the probabilities are estimated from the cell values so that Shannon's information measure can be applied, one for the output source $H(Z)$, one for the input source $H(A)$, and one for the joint events $H(A,Z)$. Then the <u>transmitted information</u> $T(A;Z)$ is defined between the input source A and the output source Z as

$$T(A;Z) = H(A) + H(Z) - H(A,Z).$$

The term $T(A;Z)$ is a measure of the dependency between the output and the input sources, and it varies from zero to a maximum value of $H(Z)$. If it is zero, then the output source is independent of the input source; that is, any knowledge of the input symbols does not help to predict the output symbols. If $T(A;Z) = H(Z)$, then the output source is completely determined by the input source; that is, every output symbol is uniquely determined by an input symbol. Another equivalent relationship for the transmitted information term is

$$T(A;Z) = H(Z) - H_A(Z),$$

where $H_A(Z) = H(A,Z) - H(A)$ is the noise term. Hence the case of a complete dependency between the input and the output sources (i.e., when $T(A;Z) = H(Z)$) is equivalent to a noiseless communication system, and the

case of no dependency between the input and the output sources (i.e., when $T(A;Z) = 0$) is equivalent to a communication system that is a noise generator.

The usual case is, however, for $T(A;Z)$ to be between zero and $H(Z)$, which reflects some dependency between the input and the output sources. Fortunately there is a test of significance by chi-square values for $T(A;Z)$ for the hypothesis of no dependency. That is, given a certain probability (either .01 or .001) of rejecting the hypothesis of no dependency, we can determine whether the obtained $T(A;Z)$ is significant (when p is chosen as .01) or "highly" significant (when p is chosen as .001). The significant and the "highly" significant T-terms imply that the output symbols can be predicted from the input symbols better than chance.

From the 1-input source and 1-output source model, we then move to the 2-input source and 1-output source model (3-dimensional contingency tables) and finally to the 3-input source and 1-output source model (4-dimensional contingency tables). The concepts involved are, however, similar to those of the 1-input source and 1-output source model. Since there is now more than one input source, there is more than one T-term, each of which measures the dependency between the input and the output sources. The chi-square distribution provides a test of the significance for each of these terms. For the 3-input sources A, B and C and the 1-output source Z, $T(ABC;Z)$ is the transmitted information between the joint input symbols from A, B and C and the output symbols from Z. When we reduce the 4-dimensional contingency tables to 3-dimensional contingency tables by summing over one input source, say, source C, then $T(AB;Z)$ is the transmitted information between the joint input symbols

from A and B and the output symbol from Z. For the present experiment, T(AB;Z) corresponds to the transmitted information between AB and Z when the responses by the male and the female Ss are pooled together. When we treat the male responses and the female responses separately, we have the T(AB;Z)/M and T(AB;Z)/F respectively. Then by further reducing the contingency tables to 2-dimensions, we have the terms T(A;Z), T(B;Z) and T(C;Z) which are the T-terms between one input source (summed over the other two sources) and the output source Z.

In the 3-input source and 1-output source, besides the T-terms, there are four 3-way interaction terms and one 4-way interaction term. The 3-way interaction terms measure the amount of transmitted information that is gained (when positive) or lost (when negative) between any two sources when the third source is known. Likewise for the 4-way interaction term, except now it has a fourth source.

There is a very close relationship between the information transmission model and that of analysis of variance. The interested reader is referred to Garner and McGill [1956]. In the remaining part of this section, we will define the information transmission model for a 3-input source and 1-output source and define formally such terms as "information", "transmitted information" and "interaction". The formulae thus derived were used in a computer program, which was written in the MAD language, to take as inputs the numbers in contingency tables and print out the different T-terms, along with their chi-square values and the interaction terms.

The information transmission model may best be illustrated by considering only one input source A and one output source Z. Let the input have $\bar{\bar{A}}$ symbols, with probability p(a) that the a^{th} input symbol

is sent, where $\sum_{a=1}^{\bar{\bar{A}}} p(a) = \sum_a p(a) = 1$; let the output have $\bar{\bar{Z}}$ symbols, with probability $p(z)$ that the z^{th} output symbol is received, where $\sum_{z=1}^{\bar{\bar{Z}}} p(z) = \sum_z p(z) = 1$; and let the joint probability that the a^{th} input symbol is sent with the z^{th} output symbol received be $p(a,z)$, where $\sum_z p(a,z) = p(a)$, $\sum_a p(a,z) = p(z)$ and $\sum_{a,z} p(a,z) = 1$. Hence we can use Shannon's information measure as follows:

$$H(A) = -\sum_a p(a) \log p(a)$$

$$H(Z) = -\sum_z p(z) \log p(z)$$

$$H(A,Z) = -\sum_{a,z} p(a,z) \log p(a,z).$$

The logarithm will be taken to the base two; hence the units for the H-terms will be in bits.

If the input and the output sources are independent (i.e., knowledge of the input symbol does not change the probability distribution of the output symbol) then $p(a,z) = p(a)\ p(z)$; hence,

$$\begin{aligned} H(A,Z) &= -\sum_{a,z} p(a,z) \log p(a,z) \\ &= -\sum_{a,z} p(a,z) \log [p(a)\ p(z)] \\ &= -\sum_{a,z} p(a,z) [\log p(a) + \log p(z)] \\ &= -\sum_a \sum_z p(a,z) \log p(a) - \sum_z \sum_a p(a,z) \log p(z) \\ &= -\sum_a p(a) \log p(a) - \sum_z p(z) \log p(z) \\ &= H(A) + H(Z). \end{aligned}$$

However if the output symbol is uniquely determined by the input symbol, then the joint probability is equal to the input source probability, i.e., $p(a,z) = p(a)$. Hence $H(A,Z)$ reduces to $H(A)$. In general, whenever the input and the output sources are not independent, then $H(A,Z) < H(A) + H(Z)$.

We define the <u>transmitted information</u> $T(A;Z)$ between the input source A and the output source Z as,

$$T(A;Z) = H(A) + H(Z) - H(A,Z).$$

$T(A;Z)$ is a term which measures the dependency or predictability of output symbols from the input symbols. It varies from zero for no dependency, to a maximum value of $H(Z)$ for complete dependency. From probability theory we know that the joint probability $p(a,z)$ is related to the conditional probability $p_a(z)$ and $p(a)$ by the equation,

$$p(a,z) = p(a)\ p_a(z).$$

Therefore we define a conditional information term $H_A(Z)$ as,

$$\begin{aligned} H_A(Z) &= -\sum_{a,z} p(a,z) \log p_a(z) \\ &= -\sum_{a,z} p(a,z) \log \frac{p(a,z)}{p(a)} \\ &= -\sum_{a,z} [p(a,z) \log p(a,z) - p(a,z) \log p(a)] \\ &= -\sum_{a,z} p(a,z) \log p(a,z) + \sum_{a} p(a) \log p(a) \\ &= H(A,Z) - H(A). \end{aligned}$$

$H_A(Z)$ denotes the information or uncertainty of the output symbol when the input symbol is known; that is, it is a measure of the noise that is generated in an information transmission model or communication system. For the present experiment, the factors which contributed to the non-uniform responses from the Ss make up the noise component. When we substitute this term into the equation for $T(A;Z)$, then we obtain the following relation:

$$T(A;Z) = H(Z) - H_A(Z).$$

In other words, the transmitted information is the difference between the information in the output source and the noise component. Thus the case of no dependency between the input and the output sources, i.e., when $T(A;Z) = 0$, is equivalent to a communication system which generates noise (i.e., a random generator). The case of complete dependency between the input and the output sources, i.e., when $T(A;Z) = H(Z)$, is equivalent to a noiseless communication system (i.e., one where each output symbol is uniquely determined by an input symbol).

In an experimental situation, these probabilities $p(a)$, $p(z)$ and $p(a,z)$ can be estimated from the observed data, in the form of contingency tables. In this 2-dimensional case, the tables are matrices with $\bar{\bar{A}}$ rows and $\bar{\bar{Z}}$ columns. Each cell (a,z) contains the number $n(a,z)$ which corresponds to the event when the a^{th} input symbol is sent out and z^{th} output symbol received. Then the number of observations of the a^{th} input symbol, $n(a)$, regardless of what is received, is $n(a) = \sum_z n(a,z)$; i.e., the row sums. Similarly the number of observations of the z^{th} output symbol, $n(z)$, regardless of what symbol is sent out is $n(z) = \sum_a n(a,z)$; i.e., the column sums. Similarly,

$$n = \sum_z n(z) = \sum_a n(a) = \sum_{a,z} n(a,z).$$

The maximum likelihood estimates of $p(a)$, $p(z)$ and $p(a,z)$ [Mood, 1950, chapter 8] are,

$$p(a) = n(a)/n$$

$$p(z) = n(z)/n$$

$$p(a,z) = n(a,z)/n.$$

Based on this estimate of the probability terms, we can calculate the various information terms.

In order to facilitate the manipulation of the estimated probabilities, let us introduce the following notation:

$$S_{az} = \frac{1}{n} \sum_{a,z} n(a,z) \log n(a,z)$$

$$S_a = \frac{1}{n} \sum_a n(a) \log n(a)$$

$$S_z = \frac{1}{n} \sum_z n(z) \log n(z)$$

$$S = \log n$$

Then,

$$H(A) = -\sum_a p(a) \log p(a)$$

$$= -\sum_a \frac{n(a)}{n} \log \frac{n(a)}{n}$$

$$= -\frac{1}{n} \sum_a n(a) \log n(a) + \frac{1}{n} \sum_a n(a) \log n$$

$$= -\frac{1}{n} \sum_a n(a) \log n(a) + \log n$$

$$= S - S_a.$$

Similarly,

$$H(Z) = S = S_z$$

$$H(A,Z) = S - S_{az}.$$

Therefore,

$$T(A;Z) = H(A) + H(Z) - H(A,Z)$$

$$= S - S_a - S_z + S_{az}.$$

Now suppose there are two input sources A and B with $\bar{\bar{A}}$ and $\bar{\bar{B}}$ symbols respectively and an output source Z with $\bar{\bar{Z}}$ symbols. Let the a^{th} input symbol from source A have a probability $p(a)$ of occurring where $\sum_{a=1}^{\bar{\bar{A}}} p(a) = \sum_a p(a) = 1$; let the b^{th} input symbol from source B have a probability $p(b)$ of occurring where $\sum_{b=1}^{\bar{\bar{B}}} p(b) = \sum_b p(b) = 1$; and let the z^{th} output symbol have a probability $p(z)$ of occurring where $\sum_{z=1}^{\bar{\bar{Z}}} p(z) = \sum_z p(z) = 1$. The information measures, in addition to those already defined, are:

$$H(B) = -\sum_b p(b) \log p(b)$$

$$H(A,B) = -\sum_{a,b} p(a,b) \log p(a,b)$$

$$H(B,Z) = -\sum_{b,z} p(b,z) \log p(b,z)$$

$$H(A,B,Z) = -\sum_{a,b,z} p(a,b,z) \log p(a,b,z)$$

$$H_{AB}(Z) = -\sum_{a,b,z} p(a,b,z) \log p_{ab}(z) = H(A,B,Z) - H(A,B).$$

For the transmitted information term, we are now considering the input as two separate sources; hence,

$$T(AB;Z) = H(A,B) + H(Z) - H(A,B,Z)$$

$$= S - S_{ab} - S_z + S_{abz},$$

where

$$S_{abz} = \frac{1}{n} \sum_{a,b,z} n(a,b,z) \log n(a,b,z)$$

$$S_{ab} = \frac{1}{n} \sum_{a,b} n(a,b) \log n(a,b).$$

Equivalently, let $H_{AB}(Z)$ be the noise generated in a 2-input and 1-output information transmission model, then

$$T(AB;Z) = H(Z) - H_{AB}(Z).$$

The number $n(a,b,z)$ is the number of observations in cell (a,b,z) of a 3-dimensional contingency table and it corresponds to the event when the a^{th} and b^{th} symbols from the respective A and B input sources are sent out and the z^{th} output symbol is received, out of a total of n observations. The number $n(a,b)$ is the number of observations in cell (a,b) summed over the Z-dimension and it corresponds to the event when the a^{th} and b^{th} symbols are sent out regardless of which symbol is received. Other S-terms for the 3-dimensional contingency tables will appear later and are assumed to be defined in similar ways.

Of the three sources, one can eliminate any one of them to consider transmitted information between the other two sources. Suppose we eliminate the source B, and consider the term T(A;Z). There are two ways in which the source B may be eliminated. One way is to sum over the B-dimension of the 3-dimensional contingency table, effectively reducing it to an A by Z 2-dimensional table, i.e., $n(a,z) = \sum_b n(a,b,z)$.

Hence,

$$T(A;Z) = H(A) + H(Z) - H(A,Z)$$
$$= S - S_a - S_z + S_{az}.$$

The other way is to compute the information transmission between A and Z separately for each symbol b from source B and then average these together. Hence,

$$T_B(A;Z) = \sum_b \frac{n(b)}{n} [T_b(A;Z)]$$
$$= \sum_b \frac{n(b)}{n} \left[- \sum_a \frac{n(a,b)}{n(b)} \log \frac{n(a,b)}{n(b)} - \sum_z \frac{n(b,z)}{n(b)} \log \frac{n(b,z)}{n(z)} + \sum_{a,z} \frac{n(a,b,z)}{n(b)} \log \frac{n(a,b,z)}{n(b)} \right]$$
$$= - \frac{1}{n} \sum_{a,b} n(a,b) \log \frac{n(a,b)}{n(b)} - \frac{1}{n} \sum_{b,z} n(b,z) \log \frac{n(b,z)}{n(b)} + \frac{1}{n} \sum_{a,b,z} n(a,b,z) \log \frac{n(a,b,z)}{n(b)}$$
$$= S_b - S_{ab} + S_b - S_{bz} - S_b + S_{abz}$$
$$= S_b - S_{ab} - S_{bz} + S_{abz}.$$

Similarly, if we were to eliminate the source A instead of B, we have

$$T(B;Z) = S - S_b - S_z + S_{bz}$$

$$T_A(B;Z) = S_a - S_{ab} - S_{az} + S_{abz}.$$

We could equally eliminate the output source Z entirely and consider the transmission between A and B;[1] hence,

---

[1] It is conventional to speak of transmission from input sources to output sources. However, the underlying contingency table for a communication system does not distinguish between input and output sources, since each source is a dimension in the table. Hence we can eliminate any one or more sources by summing over the corresponding dimensions in the table. When the output source Z is eliminated in this case, we could consider temporarily, either source A or source B as the output source. Because, by design, we have all combinations of the four grammatical forms for both the first clause (source A) and the second clause (source B), $T(A;B) = 0$.

$$T(A;B) = S - S_a - S_b + S_{ab}$$

$$T_Z(A;B) = S_z - S_{az} - S_{bz} + S_{abz}.$$

The difference between the two ways of eliminating any once source is defined as the 3-way <u>interaction information</u> $A(ABZ)$. It measures the amount of transmitted information that is gained (when positive) or lost (when negative) between any two sources when the third source is also known. If $A(ABZ)$ is zero, it implies that one of the three sources is superfluous, and therefore the 3-dimensional contingency table may be reduced to a 2-dimensional contingency table without any loss of generality. It turns out that this interaction term is the same no matter which particular source is eliminated. Thus,

$$\begin{aligned} A(ABZ) &= T_B(A;Z) - T(A;Z) \\ &= T_A(B;Z) - T(B;Z) \\ &= T_Z(A;B) - T(A;B) \\ &= -S + S_a + S_b + S_z - S_{ab} - S_{az} - S_{bz} + S_{abz}. \end{aligned}$$

Finally we can relate the 3-dimensional information transmission $T(AB;Z)$ with the 2-dimensional information transmission terms, as follows:

$$\begin{aligned} T(AB;Z) &= T(A;Z) + T(B;Z) + A(ABZ) \\ &= T_B(A;Z) + T_A(B;Z) - A(ABZ). \end{aligned}$$

The Venn diagram in Figure 4-3 also shows the relationship when $A(ABZ)$ is positive.

As our experimental results are best analyzed one frame at a time by a 3-input source and 1-output source model, we need to extend the calculations to include a third input source C. It has $\bar{\bar{C}}$ symbols, each with a probability $p(c)$ of occurring with $\sum_{c=1}^{\bar{\bar{C}}} p(c) = \sum_c p(c) = 1$. The data will be in 4-dimensional contingency tables where the cell (a,b,c,z)

has a number $n(a,b,c,z)$ which corresponds to the number of observations of the event that the a^{th}, b^{th} and c^{th} symbols from input sources A, B and C respectively are sent out with the z^{th} output symbol received.

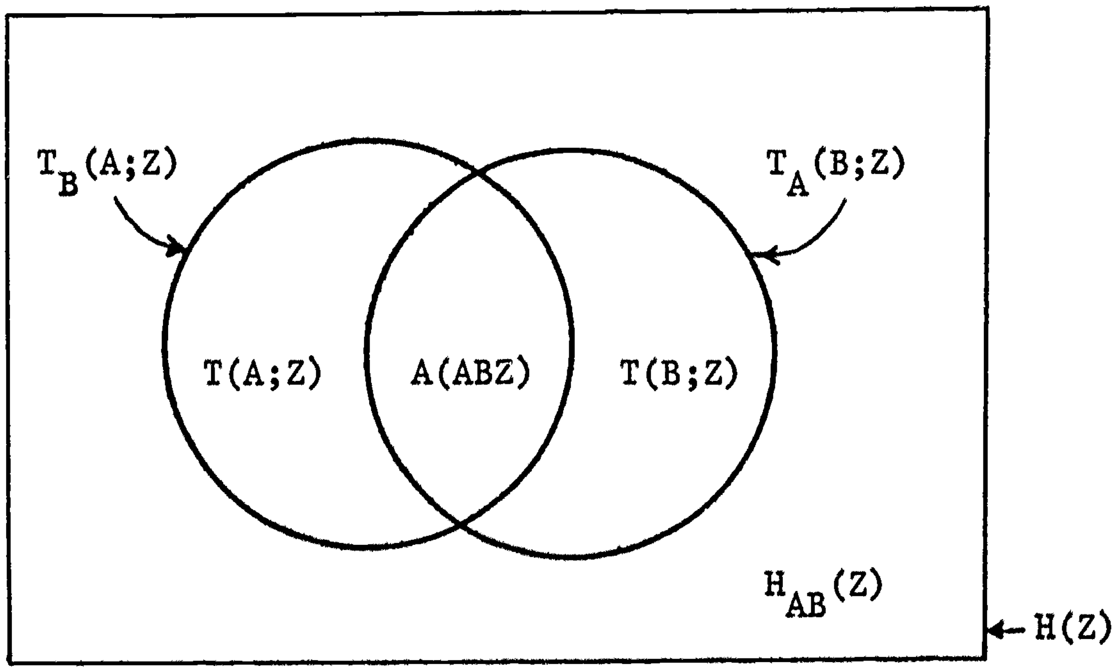

Figure 4-3: Components of transmitted information in a 2-input and 1-output model.

When the terms $n(a,b,c,z)$ are summed over any dimension, e.g., dimension A, then the n-terms are denoted as

$$n(b,c,z) = \sum_a n(a,b,c,z).$$

In general, the summed terms are denoted as follows,

$$n = \sum_a n(a) = \sum_{a,b} n(a,b) = \sum_{a,b,c} n(a,b,c) = \sum_{a,b,c,z} n(a,b,c,z).$$

The S-terms are defined in a similar way as before, e.g.,

$$S_{abcz} = \frac{1}{n} \sum_{a,b,c,z} n(a,b,c,z) \log n(a,b,c,z)$$

$$S_{abc} = \frac{1}{n} \sum_{a,b,c} n(a,b,c) \log n(a,b,c).$$

The H-terms are also defined in similar ways as before, e.g.,

$$H(A,B,C,Z) = - \sum_{a,b,c,z} p(a,b,c,z) \log p(a,b,c,z)$$

$$H(A,B,C) = -\sum_{a,b,c} p(a,b,c) \log p(a,b,c)$$

$$H_{ABC}(Z) = -\sum_{a,b,c,z} p(a,b,c,z) \log p_{abc}(z)$$

$$= H(A,B,C,Z) - H(A,B,C).$$

For this 3-input and 1-output information transmission model, we define the transmitted information $T(ABC;Z)$ between the three input sources and one output source as follows,

$$T(ABC;Z) = H(A,B,C) + H(Z) - H(A,B,C,Z)$$

$$= S - S_{abc} - S_z + S_{abcz}.$$

Equivalently, in terms of the noise term $H_{ABC}(Z)$,

$$T(ABC;Z) = H(Z) - H_{ABC}(Z).$$

There are now four 3-way interaction terms, $A(ABZ)$, $A(ACZ)$, $A(BCZ)$ and $A(ABC)$, and a 4-way interaction term $A(ABCZ)$. This 4-way interaction term denotes the amount of transmitted information gained (or lost) in transmission by controlling a fourth source when any three of the other sources are known. Some of the relationships of the 4-dimensional transmitted information to the 3-dimensional and 2-dimensional transmitted information terms are as follows,

$$T(ABC;Z) = T(A;Z) + T(B;Z) + T(C;Z) + A(ABZ) + A(ACZ) + A(BCZ) + A(ABCZ)$$

$$= T(AB;Z) + T(C;Z) + A(ACZ) + A(BCZ) + A(ABCZ)$$

$$= T_C(AB;Z) + T(C;Z)$$

$$= T(A;Z) + T(B;Z) + T_{AB}(C;Z) + A(ABZ).$$

The Venn diagram in Figure 4-4 also depicts the various components that make up the output information $H(Z)$.

For each of the T-terms we can derive the large sample distribution by the likelihood ratio test. Suppose in the 3-input and 1-output information transmission model, we want to test the hypothesis that the output source is independent of the input sources, i.e.,

$$p(a,b,c,z) = p(a,b,c)\, p(z).$$

These probabilities are estimated from sample data in the form of contingency tables by the maximum likelihood estimator:

$$p(a,b,c,z) = n(a,b,c,z)/n$$

$$p(a,b,c) = n(a,b,c)/n$$

$$p(z) = n(z)/n.$$

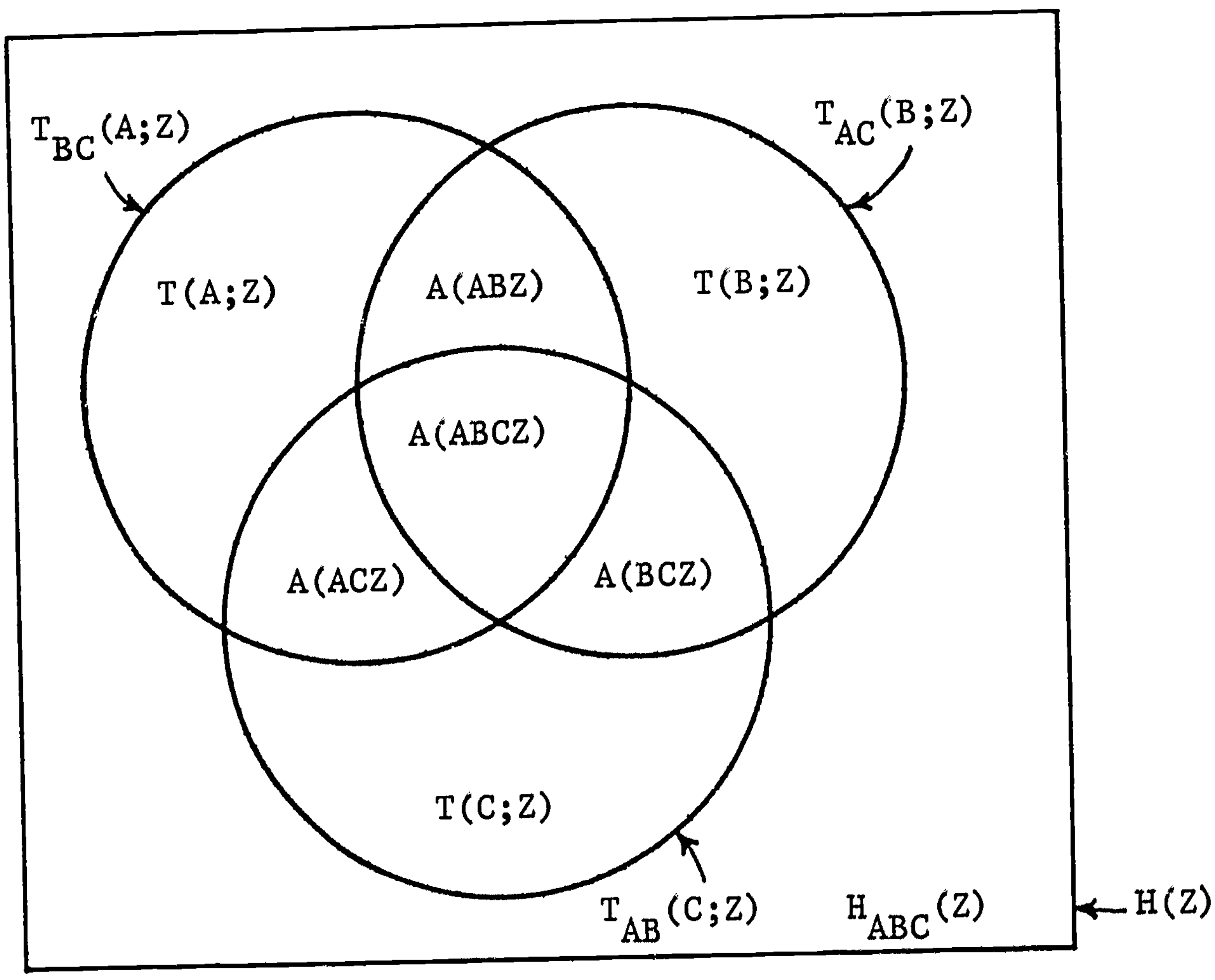

Figure 4-4: Components of transmitted information in a 3-input and 1-output model.

Therefore the independence of input and output sources can be tested by the likelihood ratio λ [Wilks, 1962, Chapter 13; Mood, 1950, chapter 12]:

$$\lambda = \frac{\prod_{a,b,c} (n(a,b,c)/n)^{n(a,b,c)} \prod_{z}(n(z)/n)^{n(z)}}{\prod_{a,b,c,z} (n(a,b,c,z)/n)^{n(a,b,c,z)}}$$

$$= \frac{n^{-2n} \prod_{a,b,c} n(a,b,c)^{n(a,b,c)} \prod_{z} n(z)^{n(z)}}{n^{-n} \prod_{a,b,c,z} n(a,b,c,z)^{n(a,b,c,z)}}$$

If we take the logarithm to the base two on both sides, we obtain,

$$\log \lambda = -2n \log n + \sum_{a,b,c} n(a,b,c) \log n(a,b,c) + \sum_{z} n(z) \log n(z)$$

$$- \left[n \log n + \sum_{a,b,c,z} n(a,b,c,z) \log n(a,b,c,z) \right]$$

$$= n \left[-\log n + \frac{1}{n} \sum_{a,b,c} n(a,b,c) \log n(a,b,c) + \frac{1}{n} \sum_{z} n(z) \log n(z) \right.$$

$$\left. - \frac{1}{n} \sum_{a,b,c,z} n(a,b,c,z) \log n(a,b,c,z) \right]$$

$$= n[-S + S_{abc} + S_z - S_{abcz}]$$

$$= -n\ T(ABC;Z).$$

From the theory of large sample distributions, it is known [e.g., Wilks, 1960, Chapter 13, p. 419] that the quantity $-2 \ln \lambda$ has an asymptotic chi-square distribution with $(\overline{\overline{ABC}} - 1)\ (\overline{\overline{Z}} - 1)$ degrees of freedom, where ln is the natural logarithm and $\overline{\overline{A}}$, $\overline{\overline{B}}$, $\overline{\overline{C}}$ and $\overline{\overline{Z}}$ are the number of symbols in the sources A, B, C and Z respectively. Therefore, by a change of base for the logarithms, where $2 \ln 2 = 1.3863$, we obtain the final expression:

$$-2 \ln \lambda = -2 \ln 2 \log \lambda$$

$$= 1.3863\ n\ T(ABC;Z).$$

Similar expressions may be derived for other T-terms. Thus we obtain the χ^2 value for any T-term by multiplying the T-term by $1.3863\ n$ where n is the sum of all the terms in a contingency table or the total number of responses for a given sentence frame and a group of <u>S</u>s. Therefore if the hypothesis

$$p(a,b,c,z) = p(a,b,c)\ p(z)$$

is true, then 1.3863 n T(ABC;Z) is distributed as $\chi^2(32)$ (chi-square for 32 degrees of freedom). For a given choice of probability level α, e.g., .01 or .001, we can find the critical points x (for p = .01) and x' (for p = .001) such that if the quantity (1.3863 n T(ABC;Z)) > x, then the hypothesis is unreasonable, or if (1.3863 n T(ABC;Z)) > x', the hypothesis is "highly" unreasonable. When the hypothesis is (or is "highly") unreasonable, we can suspect some dependency between the input and output sources, or more informally, we suspect that the Ss are responding to the cues in the sentences so as to produce a more uniform choice of the pronominal referent.

Chapter 5

RESULTS

This chapter describes the results of the experiments with the four age groups, from the college sophomore group to the fifth grade group. For each age group, the results are analyzed by the proposed three mathematical models: the Bernoulli trials model, the k-limited transducer model and the information transmission model; the presentation of the results is followed by an informal discussion which summarizes the findings.

5.1 The College Sophomore Group

The basic data for the college sophomore group appear in Tables 5-1, 5-2, and 5-3. There were 21 male Ss with an average age of 19.6 years and 20 female Ss also with an average age of 19.6 years. They were students at the University of Michigan and served as paid Ss. Table 5-1 contains the results of three sentence frames of the sentence type Each other; Table 5-2 contains the results of four sentence frames of the sentence type He Xed him back; and Table 5-3 contains the results of four sentence frames of the sentence type He Xed him. The results are tabulated separately for the male and the female Ss and their pooled data are denoted as "Total". The difference between the male and female responses for each sentence frame is calculated as a chi-square value. The $\chi^2(16)$ (chi-square values for 16 degrees of freedom) for the different sentence frames appear in the last column (χ^2_{sex}) in Table 5-5. An asterisk next to the $\chi^2(16)$ denotes that it has exceeded the .01 significance level. Each

4 x 4 matrix represents the 16 possible combinations of sentences of a frame where the rows indicate the grammatical forms (A, $\overline{A}$, $\overline{P}$ or P) of the first clause and the columns indicate the grammatical forms (A, $\overline{A}$, $\overline{P}$ or P) of the second clause. The number in each cell is the number of Ss who responded with the grammatical object-slot for that sentence. For example, in Table 5-1, the top-left matrix is the response matrix for 21 male Ss to the nonsense Each other frame. The cell A-A corresponds to the sentence John and Bill Yed each other and he Xed him with Y and X being nonsense verbs; its cell number 4 indicates that four male Ss responded with the grammatical object-slot (Bill) as the referent of he. Hence 21 - 4 = 17 male Ss responded with the grammatical subject-slot (John) as the referent of he. (Please note that for ease of usage, the first-mentioned person (John) is defined as the subject while the second-mentioned person (Bill) is defined as the object).

a. Analysis by the Bernoulli trials model. The maximum likelihood estimator $\hat{p}$ for each sentence frame (i.e., a matrix) appears at the bottom of the matrices. Next to the value of $\hat{p}$ is the confidence interval [u,v] where the tail probability α is chosen as .01. Hence by assuming a Bernoulli trials model, i.e., all the n Ss selected independently, and with the same probability $\hat{p}$ of selecting the grammatical object-slot as the referent of the pronoun, the probability of obtaining Y object-slots is $< .01$, if $Y \leq u$ or $Y \geq v$. Those cells whose numbers are $\leq u$ or $\geq v$ are starred. The sentences corresponding to these cells are interpreted as the unambiguous sentences. Since all the $\chi^2(16)$ values for the differences between the male and the female responses are not significant, subsequent discussion will pertain to the "Total" matrices, i.e., the pooled data.

First, all the sentences of the type Each other in Table 5-1 are not starred, and hence they are interpreted as ambiguous, which is what we had expected. Apparently whenever the Ss encountered sentences of the Each other type, they realized the impossibility of deciding the referent, hence approximately three quarters of them just marked an "x" on top of the first word (the first word is also the first-mentioned person, which, for the present study, is denoted as the subject). The grammatical forms of the two clauses and the different verb-pairs did not alter the predominantly subject-slot responses. This bias toward the subject-slot is reflected by the very low value for the maximum likelihood estimator $\hat{p}$; and as a consequence, the confidence interval would cover all the 16 responses in each sentence frame.

Even in the extreme, if all except one or two Ss (out of 41) selected the subject-slot response (i.e., the first word) in all 16 sentences of a frame, it might appear that all 16 sentences should be interpreted as unambiguous. But such biased responses would lead to an estimated $\hat{p}$ of very small value, and a 98% confidence interval that is certain to include one and two. Thus according to the Bernoulli trials model, all these sentences would be considered as ambiguous, since their responses can be predicted without regard to the structure (i.e., the grammatical forms of the two clauses) of the sentence. Otherwise we are forced to consider both the following two sentences of different grammatical structure as having the same unambiguous interpretation, i.e., he = John:

John and Bill Yed each other and he Xed him. [A-A]
John and Bill Yed each other and he was Xed by him. [A-P]

But this leads to an explicit contradiction, i.e., if "he = John" is accepted as the correct interpretation for one sentence, then "he = John" cannot be accepted as the correct interpretation for the other sentence,

Table 5-1

Response matrices for 21 male and 20 female college sophomore subjects

Sentence type: John and Bill Yed each other and he Xed him.

Male

Nonsense Verbs	A	$\bar{A}$	$\bar{P}$	P
A	4	7	7	4
$\bar{A}$	4	5	10	8
$\bar{P}$	5	5	8	6
P	5	7	7	9

$\hat{p}$ = .30; [2,12]

Similar Verbs	A	$\bar{A}$	$\bar{P}$	P
A	6	6	6	6
$\bar{A}$	4	5	14*	8
$\bar{P}$	6	5	10	10
P	5	6	8	8

$\hat{p}$ = .34; [2,13]

Logical Reversal Verbs	A	$\bar{A}$	$\bar{P}$	P
A	7	4	10	10
$\bar{A}$	7	7	8	8
$\bar{P}$	7	6	9	10
P	5	8	10	9

$\hat{p}$ = .37; [3,14]

Female

Nonsense Verbs	A	$\bar{A}$	$\bar{P}$	P
A	4	3	8	4
$\bar{A}$	5	5	3	3
$\bar{P}$	6	3	4	3
P	3	4	3	5

$\hat{p}$ = .21; [1,10]

Similar Verbs	A	$\bar{A}$	$\bar{P}$	P
A	5	5	5	4
$\bar{A}$	4	3	4	6
$\bar{P}$	4	4	3	5
P	3	4	5	4

$\hat{p}$ = .21; [1,10]

Logical Reversal Verbs	A	$\bar{A}$	$\bar{P}$	P
A	3	4	4	5
$\bar{A}$	5	4	3	3
$\bar{P}$	5	1	4	4
P	2	4	4	5

$\hat{p}$ = .19; [0,9]

Total

Nonsense Verbs	A	$\bar{A}$	$\bar{P}$	P
A	8	10	15	8
$\bar{A}$	9	10	13	11
$\bar{P}$	11	8	12	9
P	8	11	10	14

$\hat{p}$ = .26; [5,18]

Similar Verbs	A	$\bar{A}$	$\bar{P}$	P
A	11	11	11	10
$\bar{A}$	8	8	18	14
$\bar{P}$	10	9	13	15
P	8	10	13	12

$\hat{p}$ = .28; [5,19]

Logical Reversal Verbs	A	$\bar{A}$	$\bar{P}$	P
A	10	8	14	15
$\bar{A}$	12	11	11	11
$\bar{P}$	12	7	13	14
P	7	12	14	14

$\hat{p}$ = .28; [5,19]

Starred cell (*) has $p < .01$

because the passive transformation in the second clause must interchange the referents of he and him. The explicit contradiction would be much more clear in the following pair of sentences of the same grammatical structure as before, but taken from the He Xed him back frame:

John Yed Bill and he Xed him back. [A-A]
John Yed Bill and he was Xed back by him. [A-P]

It seems reasonable to require that the set of unambiguously interpreted sentences does not contain explicit contradictions. Thus if "he = John" is the acceptable interpretation for one of these sentences, then "he = John" cannot be the acceptable interpretation for the other sentence. Similarly, the interpretation "he = Bill" cannot be acceptable for both sentences. Hence, even though the group response may be uniform, they are made without regard for the grammatical structure of the sentence, and those sentences must be classified as ambiguous.

Now for the responses from the sentence type He Xed him back, which appear in Table 5-2, almost all the sentences are unambiguously interpreted. Contrary to the Each other sentence type, the grammatical forms of the two clauses largely determine the referents in this sentence type.

The word "back" was intended to switch the actors of the two actions, i.e., the subject of the first clause becomes the object of the second clause or the object of the first clause becomes the subject of the second clause. However, in a passive transformation, the grammatical subject (in the surface structure) becomes the logical object (in the deep structure), and likewise, the grammatical object becomes the logical subject. Therefore the switch of subject and object must be done after the two clauses are transformed into their respective deep structures, i.e., we are switching the logical subject and the logical object. Since the negative construction does not alter the relationship between the subject and the object,

Table 5-2

Response matrices for 21 male and 20 female college sophomore subjects.
Sentence type: <u>John Yed Bill and he Xed back.</u>

Nonsense Verbs

		A	$\bar{A}$	$\bar{P}$	P
Male	A	21*	21*	0*	4*
	$\bar{A}$	15	20*	4*	6
	$\bar{P}$	4*	7	13	12
	P	8	5*	13	15
	$\hat{p}$ = .50; [5,17]				
Female	A	20*	18*	2*	2*
	$\bar{A}$	14	19*	1*	2*
	$\bar{P}$	4*	7	10	9
	P	8	7	12	13
	$\hat{p}$ = .46; [4,15]				
Total	A	41*	39*	2*	6*
	$\bar{A}$	29*	39*	5*	8*
	$\bar{P}$	8*	14	23	21
	P	16	12*	25	28*
	$\hat{p}$ = .48; [12,28]				

Similar Verbs

		A	$\bar{A}$	$\bar{P}$	P
Male	A	20*	21*	3*	2*
	$\bar{A}$	18*	19*	6*	9
	$\bar{P}$	3*	4*	19*	11
	P	4*	8	17*	18*
	$\hat{p}$ = .54; [6,17]				
Female	A	20*	18*	3*	4*
	$\bar{A}$	17*	19*	4*	2*
	$\bar{P}$	2*	4*	17*	14
	P	5*	2*	17*	13
	$\hat{p}$ = .50; [5,16]				
Total	A	40*	39*	6*	6*
	$\bar{A}$	35*	38*	10*	11*
	$\bar{P}$	5*	8*	36*	25
	P	9*	10*	34*	31*
	$\hat{p}$ = .52; [14,30]				

Logical Reversal Verbs

		A	$\bar{A}$	$\bar{P}$	P
Male	A	2*	2*	8	19*
	$\bar{A}$	9	6	15*	6
	$\bar{P}$	6	19*	3*	12
	P	18*	16*	3*	2*
	$\hat{p}$ = .43; [4,15]				
Female	A	1*	1*	9	14*
	$\bar{A}$	8	0*	15*	4
	$\bar{P}$	10	20*	1*	6
	P	19*	17*	3*	2*
	$\hat{p}$ = .41; [3,14]				
Total	A	3*	3*	17	33*
	$\bar{A}$	17	6*	30*	10*
	$\bar{P}$	16	39*	4*	18
	P	37*	33*	6*	4*
	$\hat{p}$ = .42; [10,26]				

Identical Verbs

		A	$\bar{A}$	$\bar{P}$	P
Male	A	21*	19*	4*	0*
	$\bar{A}$	19*	19*	1*	4*
	$\bar{P}$	4*	5*	19*	17*
	P	8	1*	15	18*
	$\hat{p}$ = .52; [6,17]				
Female	A	20*	19*	3*	1*
	$\bar{A}$	18*	20*	2*	2*
	$\bar{P}$	3*	5*	20*	17*
	P	6*	3*	17*	18*
	$\hat{p}$ = .54; [6,17]				
Total	A	41*	38*	7*	1*
	$\bar{A}$	37*	39*	3*	6*
	$\bar{P}$	7*	10*	39*	34*
	P	14*	4*	32*	36*
	$\hat{p}$ = .53; [14,30]				

Starred cell (*) has $p < .01$

the referent of *he* is expected to be invariant under the negative transformation. Hence with respect to the matrices, we would expect the four top-left cells (A-A, A-$\bar{A}$, $\bar{A}$-A and $\bar{A}$-$\bar{A}$) to have the same referent, the four top-right cells (A-$\bar{P}$, A-P, $\bar{A}$-$\bar{P}$ and $\bar{A}$-P) to have the same referent (though not the same as the top-left ones), the four bottom-left cells ($\bar{P}$-A, $\bar{P}$-$\bar{A}$, P-A and P-$\bar{A}$) to have the same referent, and finally the four bottom-right cells ($\bar{P}$-$\bar{P}$, $\bar{P}$-P, P-$\bar{P}$ and P-P) to have the same referent. In addition, since the passive transformation effectively interchanges the subject-object relationship, its effect on the pronominal referent is expected to be the same whether the passive transformation occurs in the first or the second clause. Therefore we would expect the eight cells in the minor diagonal (i.e., the four top-right cells and the four bottom-left cells) to have the same referent. Likewise, when both clauses have the passive transformation, as in the four bottom-right cells, we would expect the two interchanges of the subject-object relationship to cancel out. Hence the referent for the eight cells in the major diagonal (i.e., the four top-left cells and the four bottom-right cells) is expected to be the same.

Except for the logical reversal frame, the response matrices of the other three sentence frames can be partitioned into four corner submatrices, each with a relatively homogeneous selection of the pronominal referent. Such partition of the total matrices confirms the expectation that the pronominal referent is invariant under the negative transformation. Further, the effect of switching the subject-object relationship by the passive transformation is confirmed by the observed "diagonal" response, i.e., the cells of the major diagonal have the same pronominal referent and the cells of the minor diagonal have the same pronominal referent (but different from the major diagonal). Hence it is reasonable to assume that the *Ss* based

their selection on the deep structure rather than the surface structure of the sentences. The fact that the cells of the major diagonal have unambiguous object-slot responses (the numbers are $\geqslant v$) and the cells of the minor diagonal have unambiguous subject-slot responses (the numbers are $\leqslant u$) supports the hypothesis that the intended meaning of "back" was indeed used by the S in selecting the referent for the pronoun he. The intended meaning was to interchange the actors in the two clauses, i.e., the logical subject of the first clause becomes the logical object of the second clause and the logical object of the first clause becomes the logical subject of the second clause.

The four top-left cells correspond to sentences with both clauses in the active forms. The pronoun he is the logical subject of the second clause, hence its expected referent would be the logical object of the first clause. Since the first clause is in the active form, its logical object is the grammatical object. The four bottom-right cells correspond to sentences with both clauses in the passive forms. Here the pronoun he is the logical object of the second clause, hence its expected referent would be the logical subject of the first clause. Since the first clause is in the passive form, its logical subject is the grammatical object. These eight cells of the major diagonal, with their cell numbers $\geqslant v$, are unambiguously interpreted in the object-slot. Thus the cells in the major diagonal are in agreement with the expected object-slot responses. The four top-right cells correspond to sentences whose first clause is in the active and the second in the passive. The pronoun he is the logical object of the second clause, hence its expected referent would be the logical subject of the first clause. Since the first clause is active, its logical subject is its grammatical subject. The four bottom-left cells correspond to sentences whose first clause is passive and its second clause active. Here the pronoun

he is the logical subject in the second clause, hence its expected referent would be the logical object of the first clause. Since the first clause is in the passive form, its logical object is its grammatical subject. These eight cells of the minor diagonal, with the cell numbers $\leqslant u$, are unambiguously interpreted in the subject-slot. Thus the cells in the minor diagonal are also in agreement with the expected subject-slot responses. There are five sentences in the nonsense frame and one sentence in the similar frame which did not attain the .01 significance level, and hence they remained as ambiguous sentences. These belong to the more syntactically complex sentences which possibly exceeded some Ss' capacity for sentence processing. These will be taken up in the following section on the k-limited transducer model.

The results for the logical reversal frame are more difficult to explain. There is still the approximate partition of the matrix into four corner submatrices, but in each submatrix, there is always one cell number which is far different from the other three cell numbers. Again the "diagonal" response is observed, save for one cell in each submatrix. Finally, the intended response reversal due to the use of logical reversal verbs is obtained, i.e., the cells of the major diagonal are unambiguously interpreted in the subject-slot while the cells of the minor diagonal are unambiguously interpreted in the object-slot. (Recall that for the other three sentence frames, the cells of the major diagonal are unambiguously interpreted in the object-slot while the cells of the minor diagonal are unambiguously interpreted in the subject-slot). The cells which differ greatly from the other three cells in each submatrix are $\overline{A}$-A, $\overline{A}$-P, $\overline{P}$-A and $\overline{P}$-P. These "deviant" cells correspond to sentences with the first clause expressed in the negative ($\overline{A}$ or $\overline{P}$) while the second clause is expressed in the affirmative (A or P). The conjunction used is always but. However, the

deviance cannot be attributed to the use of *but*, since there are three other constructions using the conjunction *but*, $A-\bar{A}$, $P-\bar{A}$ and $P-\bar{P}$, and yet their responses were unambiguous. Nor can the deviance be attributed to the particular verb-pairs used in these four cells, since the same verb-pairs, when used in other constructions, give unambiguous responses. But the fact that the sentences corresponding to these "deviant" cells are combinations of negative and affirmative clauses makes the following explanation plausible.

The set of logical reversal verb-pairs would produce the "reversal" responses if we assume that the first clause stimulated the *S*s to create a mental image of a previous action in which the second clause is a reply or consequence of that previous action. But when the first clause is negative and the second clause affirmative, it is very difficult to create a "nonoccurring" previous action such that the affirmative second clause is a reply or consequence. Hence, to some *S*s, the verb-pairs in these sentences had the same effect as if the verbs were similar. In fact, the sentence which corresponded to $\bar{A}-P$, *John did not remember Bill but he was phoned back by him*, was unambiguously interpreted in the subject-slot (*he* = *John*), which would have been the expected response if <remember, phone> did not belong to the class of logical reversal verb-pairs.

Finally we come to the responses for the *He Xed him* type, which appear in Table 5-3. The only difference between the present sentence type and the *He Xed him back* type is the lack of the word "back". But without such an explicit word to force the most likely interpretation of "to interchange the actors in the two actions", the pronominal referents of sentences in the *He Xed him* type are less determinant.

Table 5-3

Response matrices for 21 male and 20 female college sophomore subjects.
Sentence type: John Yed Bill and he Xed him.

Male

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	13	6	8	7
$\bar{A}$	9	15*	6	5
$\bar{P}$	8	8	15*	9
P	9	12	5	9

$\hat{p}$ = .43; [4,15]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	11	20*	5*	11
$\bar{A}$	16	15	4*	13
$\bar{P}$	6	5*	19*	15
P	7	4*	15	14

$\hat{p}$ = .51; [5,17]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	4	2*	9	16*
$\bar{A}$	12	1*	17*	5
$\bar{P}$	7	16*	2*	8
P	16*	15*	4	1*

$\hat{p}$ = .40; [3,15]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	20*	21*	1*	4*
$\bar{A}$	21*	21*	2*	4*
$\bar{P}$	4*	5*	19*	20*
P	5*	7*	19*	21*

$\hat{p}$ = .58; [7,18]

Female

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	18*	5	6	6
$\bar{A}$	7	12	7	1*
$\bar{P}$	6	8	11	6
P	3	5	5	12

$\hat{p}$ = .37; [3,14]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	17*	18*	1*	12
$\bar{A}$	11	14	2*	4*
$\bar{P}$	6	8	16*	14
P	10	6	15	13

$\hat{p}$ = .52; [5,16]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	2*	0*	5	15*
$\bar{A}$	10	0*	16*	1*
$\bar{P}$	6	17*	0*	8
P	18*	13*	1*	2*

$\hat{p}$ = .36; [3,13]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	20*	20*	4*	4*
$\bar{A}$	20*	20*	5*	5*
$\bar{P}$	4*	6*	17	17
P	7*	3*	20*	19*

$\hat{p}$ = .60; [7,18]

Total

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	31*	11	14	13
$\bar{A}$	16	27*	13	6*
$\bar{P}$	14	16	26*	15
P	12	17	10	21

$\hat{p}$ = .40; [9,25]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	28	38*	6*	23
$\bar{A}$	27	29*	6*	17
$\bar{P}$	12*	13*	35*	29*
P	17	10*	30*	27

$\hat{p}$ = .51; [14,29]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	6*	2*	14	31*
$\bar{A}$	22	1*	33*	6*
$\bar{P}$	13	33*	2*	16
P	34*	28*	5*	3*

$\hat{p}$ = .38; [9,24]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	40*	41*	5*	8*
$\bar{A}$	41*	41*	7*	9*
$\bar{P}$	8*	11*	36*	37*
P	12*	10*	39*	40*

$\hat{p}$ = .59; [17,32]

Starred cell (*) has $p < .01$

For the nonsense frame where the Ss could only rely on the grammatical form of the two clauses, only four of the 16 sentences attained the .01 level of being unambiguously interpreted. In fact, the responses were so non-uniform that even the expected partition of the nonsense response matrix into the four corner submatrices was not clear. The estimated $\hat{p}$ is smaller than .50, which implies that when the Ss were unable to decide on the referent, the first word (the first-mentioned person) was selected. This is clearly the same type of bias that was observed in the completely ambiguous sentence type Each other.

In the other extreme, when the verbs were from the class of identical verb-pairs, all the 16 sentences were unambiguously interpreted. Here the meaning of the verbs provided some real-world situations to help resolve the pronominal reference. The fact that the two verbs were identical probably also helped the Ss to select the most probable interpretation, i.e., as in the He Xed him back type, "to interchange the two actors in the two actions". In fact the results showed a "diagonal" response in which the cells of the major diagonal were interpreted in the object-slot while the cells of the minor diagonal were interpreted in the subject-slot. The alternative choice would result in one of the following four interpretations:

1) the two clauses of the sentence are exact repetitions of each other, e.g., in the A-A construction (or $\overline{A}$-$\overline{A}$, $\overline{P}$-$\overline{P}$, P-P),

   John Xed Bill and he (John) Xed him;

2) the two clauses are repetitions in a different grammatical form, e.g., in the A-P construction (or $\overline{A}$-$\overline{P}$, $\overline{P}$-$\overline{A}$, P-A),

   John Xed Bill and he (Bill) was Xed by him;

3) the two clauses are contradictory of each other, e.g., in the A-$\overline{\text{A}}$ construction (or $\overline{\text{A}}$-A, $\overline{\text{P}}$-P, P-$\overline{\text{P}}$),

John Xed Bill but he (John) did not X him;

4) the two clauses are contradictory of each other, but expressed in different grammatical forms, e.g., in the A-$\overline{\text{P}}$ construction (or $\overline{\text{A}}$-P, $\overline{\text{P}}$-A, P-$\overline{\text{A}}$).

John Xed Bill but he (Bill) was not Xed by him.

The first two interpretations are acceptable, emphasizing the fact that John Xed Bill, but the constructions are very awkward. The latter two interpretations give contradictory statements about "who Xed whom", if we assume John and Bill are the only two persons involved. Therefore, the lack of a reasonable alternative interpretation makes the interpretation "to interchange the two actors in the two actions" the one uniformly adopted.

For the class of similar verb-pairs, only ten of the 16 sentences were interpreted unambiguously. Because the two verbs are different, it is possible to interpret the sentences in both ways, i.e., either "the same person performed both actions" or "each person performed one action". However the latter interpretation is the more likely one, because there exists a better construction to express the fact that "the same person performed both actions", viz., John Yed and Xed Bill. In this frame, there is an overt partition of the response matrix into four corner submatrices. Under this partition, the cells of the major diagonal are interpreted in the object-slot (though the data for some of the cells did not attain the .01 level) while the cells of the minor diagonal are interpreted in the subject-slot (through the data for some did not attain the .01 level). Such an observed "diagonal" response substantiates that the interpretation "each person performed one action" was

indeed the more plausible one. It is strange, however, that the A-A and the $\bar{A}$-A sentences turned out to be ambiguous, since these are the most simple constructions. The writer cannot offer any logical explanations, except sampling error. This effect does not occur with other age groups, nor does it occur for the other sentence frames in the college sophomore group.

Finally for the logical reversal verbs, the responses were very much like those for the *He Xed him back* frame. Since this class of verbs depended on the creation of a prior action by the "inward" verbs in the first clause in which the second clause is a reply to that prior action, the word "back" did not add any more information. Again when the sentences were constructed from a negative clause ($\bar{A}$ or $\bar{P}$) followed by an affirmative clause (A or P), the causal stimulus for the creation of a prior action was absent. Hence the responses to these four sentences are either random (pronomianal referent remained ambiguous) or attained the .01 level (only cell $\bar{A}$-P), but *not* with the "reversal" interpretation. Again, except for these four "deviant" cells, the total response matrix can be partitioned into four corner submatrices, where the major diagonal is unambiguously interpreted in the subject-slot while the minor diagonal is unambiguously interpreted in the object-slot, i.e., the expected "reversal" way.

b. *Analysis by the k-limited transducer model*. Since we are dealing with one age group, this model is inteded to account for the observed degeneration (i.e., tendency toward random responding) among the 16 sentences in a frame. Since all the sentences in the *Each*

other type were responded to randomly, we cannot talk about degeneration. As for the other two sentence types, since there is a partition of every response matrix into four submatrices, each with relatively homogeneous numbers, we shall consider the degeneration in the four corners. As a rough measure of this degeneration, we shall use the number of starred cells in each corner submatrix, i.e., the number of sentences that were responded to unambiguously. Table 5-4 gives the number of starred cells in each corner submatrix for the different frames in the sentence types He Xed him and He Xed him back. The

Table 5-4

Number of starred cells in each corner submatrix
for the college sophomore subjects

	Nonsense	Similar	Log. Rev.	Identical
He Xed him	2,1,0,1	2,2,3,3	3,2,3,3	4,4,4,4
He Xed him back	4,4,2,1	4,4,4,3	3,2,3,3	4,4,4,4

4-tuple (e,f,g,h) denotes that there are e starred cells in the top-left corner (active-active submatrix), f starred cells in the top-right corner (active-passive submatrix), g starred cells in the bottom-left corner (passive-active submatrix), and h starred cells in the bottom-right corner (passive-passive submatrix). For the logical reversal frames, we do not count the "deviant" cells, hence the maximum value for each element of the 4-tuple is three.

We observe a nonincreasing relationship for the elements of the 4-tuple (e,f,g,h) across all four age groups, i.e., $e \geq f \geq g \geq h$. This relationship confirms our expectation that the passive transformation is the major contributor to the complexity in a sentence. In terms

of the k-limited transducer model, the needed k for correct decoding increases from sentences with active-active clauses to sentences with active-passive clauses to sentences with passive-active clauses to sentences with passive-passive clauses. The fact that sentences with active-passive clauses are generally less ambiguous than sentences with passive-active clauses (i.e., $f \geqslant g$), though all of the same length, is explained by including the processing as part of the k. The passive construction in the first clause requires complete decoding while the passive construction in the second clause needs only to be decoded to obtain the logical function (subject or object in the deep structure) of the pronoun <u>he</u>.

The exceptions to the nonincreasing relationship come from element f in the two logical reversal frames and the similar <u>He Xed him</u> frame, and element h from the nonsense <u>He Xed him</u> frame. For the logical reversal frames, the cell $A\text{-}\overline{P}$ was not starred, but the "deviant" $\overline{A}\text{-}P$ cell was (both belong to the active-passive submatrix), a fact which is also true for the eighth grade <u>S</u>s. Thus it appears that for the four cells in the active-passive submatrix, the sentence for cell $A\text{-}\overline{P}$ was interpreted as ambiguous while the sentence for the "deviant" cell $\overline{A}\text{-}P$ was interpreted as unambiguous, but not in the "reversal" way. Thus the lower value for element f was partly due to our exclusion of the "deviant" cells. As for the <u>He Xed him</u> frame, the lower values in elements e and f were due to the peculiar responses of the <u>S</u>s for this group, as observed in the previous section. The higher value of $h = 1$ (compared to $g = 0$) in the nonsense <u>He Xed him</u> frame is probably a sampling error caused by the particular confidence interval.

In comparing the starred cells for the two sentence types, we note that both for the identical and logical reversal frames, the elements of 4-tuples (e,f,g,h) are maximal, i.e., all have the number 4 (3 for logical reversal frames). For the nonsense and similar frames, however, even the element e (active-active combination) has not attained the maximal value in the He Xed him sentence type. This implies that the two sentence types were responded to alike (see Table 5-6 and the discussion in section 5.1.c) when the verb-pairs were from the class of identical or logical reversal verbs, and were responded to quite differently when the verb-pairs were from the class of nonsense and similar verbs.

The word "back" in the He Xed him back sentence type make the interpretation "to interchange the two actors in the two actions" so forceful that nearly all the sentences in the four frames were unambiguously interpreted. The "program" for such an interpretation is quite simple, since the only knowledge needed is the subject-object relationship; hence it is reasonable to assume that none of the decoded messages had exceeded the assignment capacity k (second transducer). But in the decoding stage (first transducer) the decoding capacity k could be exceeded when sentences contained passive constructions. Thus we see that the values of e are maximal for all four frames, reflecting the fact that as long as the sentences were decoded correctly, they were processed correctly. However, if a sentence contained passive constructions, the Ss could very easily make the mistake of transforming it to its active form without interchanging the two persons, for example, John was Yed by Bill into John Yed Bill. This type of error could easily occur because both

John and Bill possessed the same lexical features, [Roberts, 1966] or that the sentence was reversible [Slobin, 1966]. That is, given any verb, real or nonsensical, it was possible to say either John Yed Bill or Bill Yed John. In contrast, for example, given the verb "hit" and the two nouns "John" and "dog", it is possible to say John hit the dog but not The dog hit John. In particular, when the verb was nonsensical, there was no semantical criterion to guide the transformation from passive to active so that the possibility of a decoding error became greater. Indeed the 4-tuple (4,4,2,1) for the nonsense He Xed him back frame indicated a sharp degeneration as the first clause became a passive construction.

As for the He Xed him frame, the 4-tuples for the identical and logical reversal frames attained the maximal values. This indicated that sentences of these two frames were correctly decoded (within the limits of the k of the first transducer kT) and correctly processed (within the limits of the k' of the second transducer $k'T$). However when the verbs were from the classes of similar and nonsense verb-pairs, even the most simple constructions (the active-active submatrix) did not attain the maximal value, i.e., $e < 4$. This suggests that even the correctly decoded messages were not processed correctly, i.e., the k' of the second $k'T$ was exceeded due to a more complex "program" for assigning the pronomial referent. This "program" must take into account the conventional usage as wll as the relationship between the two actions, but neither criterion was usable when the verbs were nonsensical.

c. Analysis by the information transmission model. The 11 sentence frames are separately analyzed by the 3-input and 1-output infor-

mation transmission model. The results appear in Table 5-5, where the first column H(Z) is the total information or uncertainty in bits in the output source Z. The next seven columns, the T-terms, are transmitted information terms as calculated in percent of H(Z). The next four columns, the A-terms, are the interaction terms, also calculated in percent of H(Z). The last column, χ^2_{sex}, is the chi-square value with 16 degrees of freedom for the response difference by the male and female Ss in the same age group. A, B and C are input sources, where A denotes the grammatical forms of the first clause, B denotes the grammatical forms of the second clause and C denotes the sex of the S who is responding. T(ABC) is short for T(ABC;Z) and denotes the transmitted information when all three input sources are known. T(AB) is short for T(AB;Z) and denotes the transmitted information when the A and B sources are known but summed over the dimension C. T(AB)/M and T(AB)/F are transmitted information T(AB;Z) calculated respectively for only the male and female responses. T(A), T(B) and T(C) are short for T(A;Z), T(B;Z) and T(C;Z), respectively, and they denote the transmitted information when only one of the input sources is known. For each T-term, the number of degrees of freedom is given immediately below it. The relationship of the interaction and the transmitted information terms are shown in the Venn diagram in Figure 4-4.

Let x(n) and x'(n) be two critical values such that $P[\chi^2 > x(n)] < .01$ and $P[\chi^2 > x'(n)] < .001$ for a chi-square distribution of n degrees of freedom. Since the T-terms have a limiting chi-square distribution, their χ^2 values are calculated, but are not shown in Table

Table 5-5

Components of transmitted information for 21 male and 20 female college sophomore subjects

		H(Z)	T(ABC)	T(AB)	T(AB)/M	T(AB)/F	T(A)	T(B)	T(C)	A(ABZ)	A([illegible]Z)	A(BCZ)	A(ABCZ)	χ^2_{sex}
	DF		31	15	15	15	3	3	1					16
He Xed him	Nonsense	.97	10.32√	7.92√	6.75	13.51√	.25	.59	.28	7.09	.23	.07	1.82	12.29
	Similar	1.00	22.24√	20.47√	22.58√	21.85√	1.00	.24	.02	19.23	.19	.28	1.29	8.51
	Log. Rev.	.96	33.99√	32.03√	26.80√	41.19√	.45	.81	.17	30.78	.15	.12	1.53	5.19
	Identical	.98	53.69√	52.09√	57.02√	50.13√	.14	.45	.03	51.50	.09	.17	1.31	4.07
He Xed him back	Nonsense	1.00	33.29√	31.85√	34.13√	32.19√	.70	4.58√	.10	26.57	.10	.20	1.05	4.71
	Similar	1.00	39.52√	37.19√	37.58√	41.35√	.64	.70	.11	35.85	.24	.10	1.89	10.49
	Log. Rev.	.98	35.19√	32.54√	29.18√	41.37√	1.10	.90	.06	30.54	.21	.40	1.98	10.95
	Identical	1.00	51.11√	49.83√	47.65√	54.65√	.04	.80	.05	48.99	.02	.11	1.11	2.25
Each other	Nonsense	.82	3.74	1.22	3.06	2.33	.03	.46	1.04*	.73	.24	.42	.82	10.35
	Similar	.85	4.68	1.79	5.09	.88	.08	.98	1.64√	.72	.22	.46	.57	14.08
	Log. Rev.	.86	5.86	1.54	2.66	1.86	.01	.73	3.59√	.80	.03	.22	.48	19.78

Checked item (√) has $p < .001$
Starred item (*) has $p < .01$

5-5. However, when the χ^2 value for a T-term exceeds $x(n)$, that T-term has an asterisk (*) next to the term, and if the χ^2 value exceeds $x'(n)$, that T-term has a check (√) next to the term. The χ^2_{sex} terms are also starred or checked according to whether their values exceed $x(16)$ or $x'(16)$ respectively.

The interpretation of these starred and checked T-terms is the following. If the input sources, either collectively or singly, do not influence the output responses, then the probability of obtaining a T-term of the magnitude shown by a starred item is less than .01, and similarly the probability would be less than .001 for a checked item. Hence, the null hypothesis that there is no influence between the input and output sources is untenable for the starred and checked items. Therefore it is safe to assume that the pronominal referent is conveyed by the input sources for these T-terms. The interpretation of a starred or checked item for χ^2_{sex} (last column) is that the responses from the male and female Ss are different. We will say that a term is highly significant when it has a check (√), i.e., the probability of obtaining such a magnitude under the null hypothesis for the term is $<.001$ ($p <.001$), or that a term is significant when it has an asterisk (*), i.e., the probability of obtaining such a magnitude under the null hypothesis is $< .01$ ($p < .01$).

The analysis by the information transmission model supplements the analysis by the Bernoulli trials model and the k-limited transducer model for each sentence frame. These two models analyze the uniformity or non-uniformity of responses for each sentence of the frame while the information transmission model analyzes the dependency between any of the input sources and the output source. This depen-

dency is quantified by the T-terms, which is a different measure of the uniformity of responses, as an average of the 16 responses in a sentence frame.

Table 5-5 shows that none of the T-terms, except T(C;Z), for the three frames of the Each other sentence type has attained the .01 significance level. This result substantiates the previous finding (by the Bernoulli trials model) that every sentence in all three sentence frames is ambiguously interpreted. That is, both the 16 sentences individually, and collectively for each frame, were responded to in a random manner, and hence they are interpreted as ambiguous. The T(C;Z) terms are starred or checked, because there is some overall bias in the male and female responses when the responses are summed over the A and B sources, i.e., when the individual sentences are not distinguishable. However, when the individual sentences are distinguishable, the male and female responses are not significantly different, as shown by the χ^2_{sex} for 16 degrees of freedom.

For the He Xed him and the He Xed him back types, the T-terms are all highly significant ($p < .001$) when either two inputs, A and B, are known or all three inputs, A, B and C, are known. In contrast, the T-terms when only one source is known are not significant (except for T(B;Z) in the nonsense He Xed him back frame, which is probably a sampling error). The interpretation of these highly significant and nonsignificant T-terms is the following: when only one input source is known, whether it be the grammatical form of the first clause (source A) or the grammatical form of the second clause (source B) or the sex of the Ss (source C), it is not possible to predict the pronominal

referent (output source Z); but if the grammatical forms of both the first and second clauses are known (i.e., the structure of the sentence is known), with or without knowing the sex of the Ss, we can predict much better than chance the pronominal referent of the he in a sentence. The fact that we can predict the pronominal referent when the structure of the sentence is known substantiates our previous finding with the Bernoulli trials models that most of the individual sentences were unambiguous. Even though only four sentences out of the 16 sentences in the nonsense He Xed him frame were unambiguous according to the Bernoulli trials model, the T(AB;Z) for the same frame is highly significant. This says that, individually, many sentences of the nonsense He Xed him frame have ambiguous pronominal referents, but collectively (since the information transmission model measures the expected or average values) the pronominal referent can be predicted much better than chance. The high interaction term, A(ABZ), means that both sources A and B are needed for producing an output response from Z. In fact, T(A;Z) and T(B;Z) are nonsignificant, which means the knowledge of the grammatical form of one of the two clauses of a sentence is not sufficient to predict the pronominal referent. By comparing the transmitted infomation T(AB)/M and T(AB)/F, there is no indication that either the male or the female Ss could perform better than the other sex.

The size of the T-terms may be a little misleading, since a T(AB;Z) of 8% implies that 92% of the total uncertainty arises from the uncontrolled noise component, and yet T(AB:Z) is highly significant. In Appendix C, Table C-1, we have computed the T(AB;Z) term for an ideal group of n Ss, all but m of whom responded uniformly in a

diagonal way, viz, n - m Ss responded with the object-slot in the major diagonal and m Ss responded with the object-slot in the minor diagonal. T(AB;Z) is a monotonically decreasing function of m. Thus for a total of 41 Ss, the 8% for T(AB;Z) corresponds to an m of 15, i.e., 15 out of 41 Ss, or 37% of the Ss, did not agree in their responses. Thus the highly significant results come from the fact that for the large sample of responses (16 x 41 = 656), a 63-37 division is significantly different from the expected 50-50 division for random responding. The highest T(AB;Z) term is 52% which corresponds to an m of 4, signifying that approximately 10% of the Ss did not agree in their responses. Since the college sophomore group is the oldest group in the present experiment, then we can expect at most 90% agreement among the 41 Ss with adult linguistic competence.

Besides determining the significance (asterisks or checks) of the T-terms, we wish to know whether the Ss based their responses on the sentences type or on the particular verb-pairs. Thus we want to find whether the responses from two classes of verb-pairs in each sentence type were different, and whether the responses from two sentence types for each class of verb-pairs were different. We can not compare the magnitudes of the T-terms, since each is calculated as an average of the responses to each frame. Hence two T-terms may be close when their responses are very different, as between the responses of the nonsense He Xed him back and the logical reversal He Xed him back frames. The T(AB;Z) terms are respectively 31.85% and 32.54%, yet from the response matrices, the major and minor diagonals are reversed with respect to each other.

A better measure is to compare the response matrices, that is as in the case of the sex difference, to calculate the chi-square differences

between the respective cells of the two matrices. Since the sex difference was found to be nonsignificant, we are now comparing the "Total" matrices. Table 5-6 tabulates the $\chi^2(16)$ values between any two sentence types for each class of verb-pairs and Table 5-7 tabulates the $\chi^2(16)$ values between any two classes of verb-pairs for each sentence type. Again the checked item indicates that its χ^2 value exceeds $x'(16)$ and the starred item indicates that its χ^2 value exceeds $x(16)$ where $P[\chi^2 > x'(16)] < .001$ and $P[\chi^2 > x(16)] < .01$. Table 5-6 reveals that college sophomores responded alike to sentence types He Xed him and He Xed him back when the verbs were identical or logical reversal. That is, for these two classes of verb-pairs, the presence or absence of the word "back" made no difference to the Ss. However when the verb-pairs were from the classes of the nonsense and the similar verbs, the word "back" made a great deal of difference. It should not be surprising that the Ss treated the Each other sentence type differently from either the He Xed him type or the He Xed him back type, no matter what the verb-pairs were.

Table 5-7 gives the $\chi^2(16)$ values between any two verb-pairs in each sentence type. The column heading NS-SI is for the comparison between the nonsense (NS) and the similar (SI) verb-pairs; the column heading NS-LR is between the nonsense and the logical reversal (LR) verb-pairs; the column heading NS-ID is between the nonsense and the identical (ID) verb-pairs, etc. As expected, there was no difference in responding for the Each other sentence type between any two classes of verb-pairs, since all were random responses. For the He Xed him back sentence type, the Ss responded alike to sentences using either nonsense or similar verb-pairs and to sentences using either similar or identical verb-pairs (but not to sentences using either nonsense

Table 5-6

Chi-square values between any two sentence types for each class of verb-pairs for the college sophomore subjects.

	Xed him vs Xed him back	Xed him vs Each other	Xed him back vs Each other
Nonsense verbs	94.5✓	57.3✓	197.8✓
Similar verbs	42.2✓	157.5✓	223.2✓
Log. Rev. verbs	9.8	144.0✓	160.8✓
Identical verbs	18.9		

Checked item (✓) has $p < .001$

Table 5-7

Chi-square values between any two verb-pairs in each sentence type for the college sophomore subjects.

	NS-SI	NS-LR	NS-ID	SI-LR	LI-ID	LR-ID
He Xed him	83.6✓	188.3✓	199.5✓	319.4✓	71.4✓	530.0✓
He Xed him back	22.3	395.3✓	41.6✓	471.2✓	16.3	534.8✓
Each other	3.8	4.3		4.8		

Checked item (✓) has $p < .001$

or identical verb-pairs). But this was not true for the He Xed him sentence type, since in the absence of the word "back", the verbs must carry the "message" about the pronominal referent. Thus each frame in the He Xed him type was responded to differently from any of the other frames. In addition, the extremely large χ^2 values for the NS-LR, SI-LR and LR-ID pairs indicated that the logical reversal verb-pairs were always treated differently from the other verb-pairs in both the He Xed him and He Xed him back types.

d. Summary. The results from the college sophomores group conclusively demonstrated that subjects could resolve an ambiguity in a sentence by reducing the meaning of a key word (here the pronoun he) from two to one. Although Kaplan [1949] demonstrated that reduction of the senses of a key word was possible through its immediate neighboring words, the present experiment demonstrated that, for some sentences, any number of neighboring words, short of the entire sentence, could not resolve the ambiguity. That is, the resolution of the ambiguous referent depended on the structure (sentence type and grammatical form) of the sentence. However, if a sentence had an inherently ambiguous structure, such as that of the Each other type, it was demonstrated that the sentence remained ambiguous with any verb-pairs and under any grammatical forms. It seems safe to generalize this finding to say that some sentences would remain ambiguous given all the available contexts in a paragraph of chapter.

The Bernoulli trials model partitioned the 16 sentences in each frame into two mutually exclusive sets: those that remained ambiguous and those that had resolved the pronominal ambiguity. With this model it was found, essentially, that all the sentences in any frame of the

Each other type were ambiguous, while nearly all the sentences in any frame of the other two types were unambiguous. Then the k-limited transducer model accounted for the degenerative responses by a sequence of two k-limited transducers: the decoding transducer (with k units) and the assignment transducer (with k' units). The degeneration, then, was due either to having exceeded the k units while decoding a sentence or to having exceeded the k' units while making the assignment for the pronoun. Finally the information transmission model computed the dependency between the input sources and the output source in terms of the transmitted information (T-terms). With the model it was found essentially that the dependency was absent for the Each other sentence type, but was highly significant when two or three input sources were known for both the He Xed him and the He Xed him back sentence types.

5.2 The Eighth Grade Group

The basic data for the eighth grade group appear in Tables 5-8, 5-9 and 5-10. There were 26 male Ss with an average of 13.5 years and an average I.Q. score of 117, and 28 female Ss with an average age also of 13.5 years and an average I.Q. score of 116. The data are pooled responses from two classes, both from the West Junior High School in Ypsilanti, Michigan. Table 5-8 contains the results of three sentence frames of the Each other type; Table 5-9 contains the results of four sentence frames of the He Xed him back type; and Table 5-10 contains the results of four sentence frames of the He Xed him type. The results are tabulated separately for the male and the female Ss, and their pooled data are denoted as "Total". The difference between the

male and female responses for each sentence frame is calculated as a chi-square value, which appears in the last column (under χ^2_{sex}) in Table 5-12. Only the similar He Xed him frame showed a significant difference ($p < .01$) between the male and the female responses.

As with the college sophomore group, each response matrix in Tables 5-8, 5-9 and 5-10 represents the 16 possible sentences in a sentence frame. The rows of each response matrix denote the grammatical forms of the second clause. Each cell of the response matrices corresponds to a particular sentence, and its cell number is the number of Ss who selected the grammatical object-slot as the referent of the pronoun he. For the nonsense frames, there were 25 male Ss instead of 26.

a. Analysis by the Bernoulli trials model. The maximum likelihood estimator $\hat{p}$ for each sentence frame (a matrix) appears at the bottom of the matrix. Next to the value of $\hat{p}$ is the confidence interval $[u,v]$ where the probability α is chosen as .01. Hence if we assume a Bernoulli trials model, i.e., all the n Ss selected the grammatical object-slot as the referent of the pronoun he independently and with the same probability $\hat{p}$, then the probability of obtaining Y object-slots is $< .01$, if $Y \leqslant u$ or $Y \geqslant v$. Those cells whose numbers are $\leqslant u$ or $\geqslant v$ are starred. The sentences corresponding to these cells are interpreted as the unambiguous sentences. Let us see if this interpretation is reasonable.

First, none of the cells of the response matrics in Table 5-8 are starred; hence all the sentences belonging to the Each other type are interpreted as ambiguous, which is what we had expected. Since the $\chi^2(16)$ for the male and female differences are all nonsignificatn, we shall dis-

Table 5-8

Response matrices for 26 male[a] and 28 female eighth grade subjects

Sentence type: John and Bill Yed each other and he Xed him.

Male

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	[illegible]	7	11	12
$\bar{A}$	11	14	14	10
$\bar{P}$	15	8	10	11
P	12	13	12	12

$\hat{p}$ = .45; [6,18]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	15	13	11	13
$\bar{A}$	9	16	13	12
$\bar{P}$	13	9	10	11
P	14	13	14	13

$\hat{p}$ = .48; [7,19]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	13	11	11	10
$\bar{A}$	10	14	12	13
$\bar{P}$	12	10	10	12
P	12	13	12	9

$\hat{p}$ = .44; [6,18]

Female

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	10	10	14	15
$\bar{A}$	13	16	13	14
$\bar{P}$	9	12	12	13
P	10	14	14	16

$\hat{p}$ = .46; [7,20]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	15	13	12	13
$\bar{A}$	13	14	9	13
$\bar{P}$	9	15	11	14
P	11	12	13	10

$\hat{p}$ = .44; [6,19]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	10	11	13	12
$\bar{A}$	9	14	12	11
$\bar{P}$	9	15	9	10
P	10	14	14	11

$\hat{p}$ = .41; [6,19]

Total

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	19	17	25	27
$\bar{A}$	24	30	27	24
$\bar{P}$	24	20	22	24
P	22	27	26	28

$\hat{p}$ = .45; [15,33]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	30	26	23	26
$\bar{A}$	22	30	22	25
$\bar{P}$	22	24	21	25
P	25	25	27	23

$\hat{p}$ = .46; [16,34]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	23	22	24	22
$\bar{A}$	19	28	24	24
$\bar{P}$	21	25	19	22
P	22	27	26	20

$\hat{p}$ = .43; [15,33]

a: Only 25 male Ss for the nonsense verbs

cuss only the pooled data. Since the estimated values of $\hat{p}$ vary from .43 to .46 (as compared to a variation from .26 to .28 for the college sophomore group), there is no marked bias toward the subject-slot. For this group there was no uniform responding when the Ss recognized the impossibility of deciding the selection of the pronominal referent. There is no question that the Ss had recognized the ambiguity in these sentences, since their responses to the other two sentence types were polarized, i.e., subject-slot responding (low numbers in the matrices) for one set of sentences and object-slot responding (high numbers in the matrices) for the remaining set of sentences.

As for the responses to the sentence type He Xed him back, which appear in Table 5-9, none of the four response matrices was completely starred. However, the responses were definitely non-random, especially when the verb-pairs were identical. In contrast to the Each other type, the word "back" in the He Xed him back type provided the cues for a most probable interpretation, i.e., "to interchange the actors of the two actions". Let us see if the responses were consistent with this expected interpretation. Again since the male and the female responses were not significantly different, only their pooled data are considered.

For the identical frame, the response matrix may be partitioned into four corner submatrices, each with relatively homogeneous numbers. Thus the responses were invariant under the negative transformation. In addition, we observe that the responses were "diagonal", that is, the cells of the major diagonal (the four top-left cells and the four bottom-right cells) have one response while the cells of the minor diagonal (the four top-right cells and the four bottom-left cells) have the other response.

Table 5-9

Response matrices for 26 male[a] and 28 female eighth grade subjects.
Sentence type: John Yed Bill and he Xed him back.

Male

	Nonsense Verbs A	$\bar{A}$	$\bar{P}$	P	Similar Verbs A	$\bar{A}$	$\bar{P}$	P	Logical Reversal Verbs A	$\bar{A}$	$\bar{P}$	P	Identical Verbs A	$\bar{A}$	$\bar{P}$	P
A	18*	19*	14	10	26*	21*	11	11	6*	5*	8	19*	25*	23*	6*	5*
$\bar{A}$	13	16	11	8	20	22*	11	12	15	5*	15	10	21*	22*	10	8*
$\bar{P}$	9	13	11	6*	12	7*	14	14	10	15	6*	11	5*	9	17	14
P	6*	11	10	8	6*	9	19	13	19*	19*	9	13	7*	9	16	22*
	$\hat{p}$ = .46; [6,18]				$\hat{p}$ = .55; [8,21]				$\hat{p}$ = .44; [6,18]				$\hat{p}$ = .53; [8,21]			

Female

	Nonsense Verbs A	$\bar{A}$	$\bar{P}$	P	Similar Verbs A	$\bar{A}$	$\bar{P}$	P	Logical Reversal Verbs A	$\bar{A}$	$\bar{P}$	P	Identical Verbs A	$\bar{A}$	$\bar{P}$	P
A	21	24*	11	8*	24*	22*	6*	9	5*	11	15	23*	27*	24*	8*	6*
$\bar{A}$	17	23*	15	15	24*	26*	9	16	20	10	17	7*	25*	25*	10*	10*
$\bar{P}$	11	11	14	13	11	9	15	14	15	18	10	17	13	12	21	20
P	17	13	17	13	11	6*	15	16	21	17	14	11	10*	6*	19	22
	$\hat{p}$ = .54; [9,22]				$\hat{p}$ = .52; [8,22]				$\hat{p}$ = .52; [8,22]				$\hat{p}$ = .58; [10,23]			

Total

	Nonsense Verbs A	$\bar{A}$	$\bar{P}$	P	Similar Verbs A	$\bar{A}$	$\bar{P}$	P	Logical Reversal Verbs A	$\bar{A}$	$\bar{P}$	P	Identical Verbs A	$\bar{A}$	$\bar{P}$	P
A	39*	43*	25	18*	50*	43*	17*	20*	11*	16*	23	42*	52*	47*	14*	11*
$\bar{A}$	30	39*	26	23	44*	48*	20*	28	35*	15*	32	17*	46*	47*	20*	18*
$\bar{P}$	20	24	25	19	23	16*	29	28	25	33	16*	28	18*	21*	38	34
P	23	24	27	21	17*	15*	34	29	40*	36*	23	24	17*	15*	35	44*
	$\hat{p}$ = .50; [18,36]				$\hat{p}$ = .54; [20,38]				$\hat{p}$ = .48; [17,35]				$\hat{p}$ = .55; [21,39]			

a: Only 25 male Ss for the nonsense verbs
Starred cell (*) has $p < .01$

This observed "diagonal" response suggests that the Ss selected the pronominal referent on the basis of the deep structure rather than the surface structure of the sentences. Further, the fact that the cells of the major diagonal were largely responded unambiguously in the object-slot and that all the cells of the minor diagonal were responded unambiguously in the subject-slot was consistent with the intended interpretation for the word "back".

There are not as many starred cells for the similar frame as for the identical frame. However, there is still a clear partition of the response matrix into four corner submatrices. Thus the responses were invariant under the negative transformation. All the starred cells form part of a "diagonal" response, which is consistent with the assumption that Ss responded to the deep structure rather than the surface structure of the sentences. Further the starred cells in the major diagonal were object-slot responses (cell number $\geqslant$ v) while the starred cells in the minor diagonal were subject-slot responses (cell number $\leqslant$ u). Thus the interpretation used for the similar frame was also "to interchange the two actors of the two actions", which is consistent with the word "back".

For the nonsense frame, however, it is meaningless to partition the matrix since most of the sentences elicited random responding. Of the four starred cells three came from the most simple constructions, i.e., when both clauses of the sentence were active (A or $\overline{A}$), while the remaining one came from the A-P construction. The next section will describe more about this degenerative behavior.

As for the logical reversal frame, we again observe the same four "deviant" cells that occurred for the college sophomore group. These are

the cells corresponding to sentences whose first clause is negative ($\overline{A}$ or $\overline{P}$) and whose second clause is affirmative (A or P). In fact, two of the four cells ($\overline{A}$-A and $\overline{A}$-P) attained the .01 significance level not in the expected "reversal" way. That is, the negative-affirmative constructions with the logical reversal verb-pairs made the creation of a "non-occurring" previous action so difficult that the Ss effectively treated the verbs as if they belonged to the class of similar verbs. Besides the four "deviant" cells, the remaining ones can be partitioned into corner submatrices. Further, the expected "reversal" response is observed, since the major diagonal has the subject-slot responding while the minor diagonal has the object-slot responding.

For the sentence type He Xed him, we have the response matrices shown in Table 5-10, Basically the results are not very different from those shown in Table 5-9 for the He Xed him back type. The responses, however, are degenerated, from nearly all unambiguous responding for the identical frame, to nearly random responding for the nonsense frame. Only the similar frame showed a significant difference between the male and the female responses. However, for ease of discussion, we shall still pool the male and the female responses, and consider the "Total" matrices.

The results from the identical frame can be partitioned into four corner submatrices. In addition, the major diagonal has the object-slot responding while the minor diagonal has the subject-slot responding. Hence, even without the word "back", the identical frame was responded to with the interpretation "to interchange the two actors of the actions".

At the other extreme, the results from the nonsense frame are nearly all random. Here the Ss lacked both the semantic cues from the verbs and

Table 5-10

Response matrices for 26 male[a] and 28 female eighth grade subjects.
Sentence type: John Yed Bill and he Xed him.

Male

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	16*	9	15	11	16	23*	8	16	1*	5*	9	19*	24*	25*	8*	9
$\bar{A}$	9	7	10	12	18	15	10	6*	12	5*	18*	7	21*	22*	11	4*
$\bar{P}$	8	5	8	6	7	5*	16	8	13	13	6	6	8*	6*	18	19
P	11	9	8	7	13	8	12	11	22*	15	11	12	5*	7*	15	16
	$\hat{p}$ = .38; [4,16]				$\hat{p}$ = .46; [6,19]				$\hat{p}$ = .42; [5,18]				$\hat{p}$ = .52; [8,20]			

Female

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	20	18	15	16	24*	20	13	20	6*	8	12	22*	26*	27*	5*	9*
$\bar{A}$	19	17	14	14	22	22	6*	14	20	9	21*	9	24*	27*	12*	10*
$\bar{P}$	12	13	11	11	17	11*	22	21	9	19	10	13	13	17	23	25*
P	11	13	20	15	15	12	19	18	22*	19	7*	10	10*	8*	22	24*
	$\hat{p}$ = .53; [9,22]				$\hat{p}$ = .62; [11,24]				$\hat{p}$ = .48; [7,21]				$\hat{p}$ = .63; [12,24]			

Total

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	36*	27	30	27	40*	43*	21*	36	7*	13*	21	41*	50*	52*	13*	18*
$\bar{A}$	28	24	24	26	40*	37	16*	20*	32	14*	39*	16*	45*	49*	23*	14*
$\bar{P}$	20	18	19	17	24	16*	38	29	22	32	16*	19	21*	23*	41*	44*
P	22	22	28	22	28	20*	31	29	44*	34*	18	22	15*	15*	37	40
	$\hat{p}$ = .46; [14,33]				$\hat{p}$ = .54; [21,39]				$\hat{p}$ = .45; [16,34]				$\hat{p}$ = .58; [23,41]			

a: Only 25 male Ss for the nonsense verbs
Starred cell (*) has $p < .01$

the structural cues from this sentence type (viz., the word "back") to assist them in giving a uniform response. The next section will attempt to characterize the degeneration by a k-limited transducer.

As for the similar frame, the partition into submatrices is still observable as well as a difference in response between the major and the minor diagonals. Again, the most plausible interpretation used by the Ss was "to interchange the two actors of the two actions", although the structural cues from the word "back" were absent. However, responses to the more syntactically complex sentences are random.

As for the logical reversal frame, the four "deviant" cells are still observable as are the four corner submatrices. However, since the more syntactically complex constructions tend to degenerate in the responses they elicit, all the cell numbers in the four bottom-right corner submatrix (sentences with passive-passive construction) are close to each other, so that it is hard to classify one of them ($\overline{P}$-P) as "deviant" from the other three. Besides these "deviant" ones, the starred cells in the major and minor diagonals confirm the expected "reversal" responding.

b. Analysis by the k-limited transducer model. Since we are dealing with one age group, this model is intended to account for the observed degeneration among the 16 sentences in each frame. Since all the sentences in the Each other type were responded to randomly, we cannot talk about degeneration. In Table 5-11 we tabulate the 4-tuples (e,f,g,h) as measures of degeneration for each frame from the He Xed him back and He Xed him types. The elements in the 4-tuples are respectively the numbers of starred cells of submatrices in the top-left corner, the top-right corner, the bottom-left corner and the bottom-right corner.

The observed fact that the elements of the 4-tuples are nonincreasing from active to passive confirms our expectation that passive transformation is the major contributor to the complexity in a sentence.

Table 5-11

Number of starred cells in each corner submatrix for the eighth grade subjects

	Nonsense	Similar	Log. Rev.	Identical
He Xed him	1,0,0,0	3,3,2,0	3,2,2,1	4,4,4,2
He Xed him back	3,1,0,0	4,3,3,0	3,1,2,1	4,4,4,1

In terms of k-limited transducers, the needed k for correct decoding increases from sentences with active-active clauses to sentences with active-passive clauses to sentences with passive-active clauses to, finally, sentences with passive-passive clauses, all without regard to the negative transformations.

For the He Xed him back type, the decoded sentences plus the assignment "programs" are within the limits of the k' of the assignment transducer for all frames except the nonsense frame. This is the case because the first elements of the 4-tuples, which correspond to sentences with active-active constructions, are maximal for these three frames (note that 3 is the maximal value for the logical reversal frame). The identical frame also has maximal values for the second and third elements of its 4-tuple, while the similar and logical reversal frames do not, which suggests that the decoding was facilitated when the two verbs were identical. This seems reasonable, since both clauses use the same verb to describe an action between two persons so that if a S had decoded one clause

correctly, then the decoding of the other clause could have been greatly facilitated. However, this interpretation would suggest that the responses to the four bottom-right cells (the passive-passive construction) should be more uniform than the four bottom-left cells (the passive-active constructions), since we would expect less interference in transferring the decoded message from one passive clause to another passive clause than in transferring from a passive clause to an active clause. In terms of degeneration and the 4-tuples (e,f,g,h), we would then expect $g < h$ for the identical frame. But we observe the contrary. The contradiction is probably due to our arbitrariness in selecting the critical point of .01 (i.e., $\alpha = .01$) in the Bernoulli trials model, but more importantly to the maximal likelihood estimator $\hat{p}$. The value of $\hat{p}$ thus computed was greater than .50 (actually .55), which is biased toward responding to the object-slots. Thus for the same critical level of α, it requires more uniform responding, as measured by percent of agreement among the Ss, to attain a significant result (i.e., being starred) when the response is for the object-slot than is for the subject-slot. Thus the Bernoulli trials model, in this case, tends to give a smaller value for h. If we take the average percent agreement among the Ss in the four corner submatrices, we would obtain the following 4-tuple: (99, 71, 67, 70). Thus, as far as the Ss' agreement in their responses is concerned, there was actually more agreement in the passive-passive constructions (fourth element) than the passive-active constructions (third element), although their numerical difference in this case is small. Hence, the result is consistent with the hypothesis that identical verb-pairs tended to facilitate the decoding of a second clause after having decoded correctly the first clause, but that the facilitation may not be enough to overcome the bias against

attaining a predetermined tail probability level.[2] Finally for the nonsense frame of the He Xed him back type, even the value of e is not maximal. Thus for the most simply constructed sentences (the active-active constructions) the eighth grade group could not make the pronominal assignment as easily for the nonsense verb-pairs as for the real verb-pairs. In terms of the k-limited transducer, we say that the "assignment program" for sentences with nonsense verb-pairs was sufficiently longer than for sentences with real verb-pairs so that the limits of the memory span k' was exceeded for some Ss. If, in addition, we introduce syntactic complexity in sentences, we observe a very quick degeneration into complete random responding, as witnessed by the 4-tuple (3,1,0,0).

For the He Xed him type, both the identical and the logical reversal frames have attained the maximal value for at least the first element of the 4-tuple (e,f,g,h), which implies that sentences with either identical or logical reversal verb-pairs were well within the limits of the k' of the assignment transducer. For these sentences, the Ss seem to do as well whether the word "back" was present or not. For the identical frame, both the second and the third elements of the 4-tuple are also maximal, but $g > h$. Again if we calculate the percent agreement for the four corner submatrices, we obtain the following 4-tuple: (91, 68, 65, 75). Thus there was actually a higher degree of agreement among the Ss' responses under passive-passive constructions than under the passive-active con-

---

[2] The data from the identical He Xed him frame in this group and from both frames in the other age groups are more convincing in supporting this hypothesis. For example, the 4-tuple in percent agreement for the identical He Xed him frame in the eight grade is (91, 68, 65, 75); the 4-tuples in percent agreement for the identical He Xed him back and He Xed him frames for the college sophomore group are (94,90, 79, 86) and (100, 82, 75, 93) respectively; the 4-tuples in percent agreement for the identical He Xed him back and He Xed him frames for the seventh grade group are (87, 69, 56, 67) and (95, 81, 60, 75) respectively.

structions. For the nonsense frames, however, the assignment was so difficult that even the correctly decoded message exceeded the limits of the k' of the assignment transducer. Likewise the non-maximal value for e in the similar frame suggests that the limits of the k' is exceeded by some Ss, and as sentences become more complex syntactically, the responses became random, as witnessed by $h = 0$.

c. Analysis by the information transmission model. The results from an analysis by the information transmission model appear in Table 5-12. The columns are components of transmitted information (the T-terms) and interaction (the A-terms) terms. Again any checked (√) term is highly significant, i.e., under the null hypothesis of no transmission between the input sources and the output source, the probability of obtaining the term of such a magnitude is $< .001$. Similarly, any starred (*) term is significant, i.e., the probability of obtaining such a magnitude under the null hypothesis is $< .01$.

The T-terms for the three frames in the Each other type are all non-significant. Thus the output response was independent from the input sources, i.e., any knowledge of the grammatical form of the first clause (input source A), the grammatical form of the second clause (input source B). and/or the sex of the S (input source C) did not influence the prediction of the pronominal referent (output source Z). This is consistent with the interpretation from the Bernoulli trials model, that each of the 16 sentences of the Each other frames elicited random responses.

For the four frames in the He Xed him back type, all the doubly-controlled (knowing input sources A and B) and triply-controlled (knowing all three input sources) T-terms are highly significant, except

Table 5-12

Components of transmitted information for 26 male[a] and 28 female eighth grade subjects. The average I.Q. scores for the male and female Ss are 117 and 116 respectively.

		H(Z)	T(ABC)	T(AB)	T(AB)/M	T(AB)/F	T(A)	T(B)	T(C)	A(ABZ)	A(ACZ)	A(BCZ)	A(ABCZ)	χ^2_{sex}
	DF		31	15	15	15	3	3	1					16
He Xed him	Nonsense	1.00	5.59√	2.43	4.20	3.48	1.79√	.26	1.78√	.39	.12	.22	1.04	26.35
	Similar	1.00	11.80√	7.91√	10.81√	9.34√	1.15*	.56	1.75√	6.20	.49	.38	1.27	32.51*
	Log. Rev.	.99	14.10√	12.27√	15.14√	12.56√	1.17*	.14	.30	10.96	.39	.11	1.03	12.61
	Identical	.98	25.48√	23.03√	23.59√	25.62√	.71	.71	.84*	21.62	.63	.11	.87	19.43
He Xed him back	Nonsense	1.00	7.65√	5.63√	6.77*	7.45√	1.53√	2.02√	.52	2.08	.57	.13	.81	15.37
	Similar	1.00	15.84√	14.36√	15.46√	16.08√	2.53√	1.15*	.06	10.68	.26	.24	.92	8.56
	Log. Rev.	1.00	10.33√	8.84√	10.58√	9.40√	.83	.33	.36	7.68	.25	.15	.72	10.20
	Identical	.99	22.46√	21.52√	22.92√	21.68√	.47	.96*	.18	20.09	.31	.14	.32	6.43
Each other	Nonsense	.99	1.88	1.20	2.13	1.66	.37	.19	.00	.63	.06	.29	.34	4.60
	Similar	1.00	1.55	.67	1.63	1.27	.14	.11	.11	.42	.16	.04	.57	5.85
	Log. Rev.	.98	1.23	.68	.83	1.45	.07	.27	.08	.35	.03	.19	.24	2.87

a: Only 25 male Ss for the nonsense frames
Checked item (√) has $p < .001$
Starred item (*) has $p < .01$

one, which is significant. Hence even though only four sentences are individually unambiguous (by the Bernoulli trials) in the nonsense frame, the knowledge of the grammatical forms of the two clauses in a sentence was sufficient to predict the response. That is, the 16 sentences of the nonsense frame, on the average, were not ambiguous. Many of the T(A) and T(B) terms are also significant or highly significant, which is caused by differences in the resulting responses when summed over the sources A and C and in the resulting responses when summed over the sources B and C. Hence there is some tendence for this younger group to "jump" to a response based on the grammatical form of the two clauses. The maximum T(AB) term comes from the identical frame. Its value is 21.5%, which corresponds to an m of 13 in Table C-1. That is, we can expect a maximum of (54 - 13)/54 = 76% agreement among the 54 Ss in their response to the He Xed him back type.

As for the He Xed him type, all the doubly-controlled and triply-controlled T-terms, except for the nonsense frame, are highly significant. Thus, on the average, sentences with real verb-pairs of this type are considered unambiguous. In other words, there were enough cues "transmitted" from the sentences to the Ss to produce uniform responding. On the other hand, when nonsense verb-pairs were used (without the word "back"), the doubly-controlled T-terms are not significant, i.e., there were not enough cues "transmitted" for uniform responding. Some of the T(A) and T(C) terms are checked or starred which is caused by some bias in the resulting responses when summed over the other two sources. For this sentence type, the maximum T(AB) is 23% which corresponds to an m value of 12 in Table C-1. Hence we can expect at most a (54 - 12)/54 = 78% agreement among the 54 eighth grade Ss in the sentence type He Xed him.

The maximal agreement for both the He Xed him and the He Xed him back types occurred when the verb-pairs were identical. The fact that the identical He Xed him has a higher agreement than the identical He Xed him back suggests some interference between the cues from the word "back" and the cues from "identical verbs". The T(AB)/M terms are not consistently greater (or less) than the corresponding T(AB)/F terms; hence there is no indication that either the male or the female Ss could out-perform the other.

One question we wish to ask is whether the Ss based their responses on the sentence type or on the particular verb-pairs. Hence we want to find out whether there is any difference between the responses from two classes of verb-pairs for the same sentence type and whether there is any difference between the responses from two sentence types with the same class of verb-pairs. In Table 5-13, we have tabulated the $\chi^2(16)$ values between any two sentence types for each class of verb-pairs. In contrast with the college sophomore group, the eighth grade Ss did not make any significant different responses between the He Xed him and the He Xed him back types. Of course the Each other type was responded to differently from the other two types. However, the nonsense He Xed him frame was not significantly different from the nonsense Each other frame; hence the nonsense He Xed him frame must also be random, a fact consistent with the finding from the Bernoulli trials model.

Table 5-14 tabulates the $\chi^2(16)$ values between any two verb-pairs in each sentence type. We note again that the Ss responded to the Each other sentence type in a way that was independent of the particular class of verb-pairs, since all the sentences elicited random responding. For the He Xed him back sentence type, the Ss responded alike to sentences

Table 5-13

Chi-square values between any two sentence types for each class of verb-pairs for the eighth grade subjects.

	Xed him vs Xed him back	Xed him vs Each other	Xed him back vs Each other
Nonsense verbs	22.7	18.6	46.9✓
Similar verbs	29.6	44.6✓	71.7✓
Log. Rev. verbs	6.3	69.1✓	54.3✓
Identical verbs	9.2		

Checked item (✓) has $p < .001$

Table 5-14

Chi-square values between any two verb-pairs in each sentence type for the eighth grade subjects.

	NS-SI	NS-LR	NS-ID	SI-LR	SI-ID	LR-ID
He Xed him	44.3✓	91.7✓	143.8✓	159.1✓	55.4✓	310.7✓
He Xed him back	30.6	118.6✓	64.7✓	199.9✓	20.2	251.9✓
Each other	7.9	5.8		2.8		

Checked item (✓) has $p < .001$

using either nonsense or similar verb-pairs and to sentences using either similar or identical verb-pairs (but not to sentences using either nonsense or identical verb-pairs). This was not true for the He Xed him sentence type, since in the absence of the cue word "back", the verbs must carry the "message" about the pronominal referent. In addition, the very large χ^2 values for the NS-LR, SI-LR and LR-ID differences in the He Xed him and the He Xed him back types indicate that the logical reversal verb-pairs were treated very differently from the other verb-pairs.

d. Summary. The results from the eighth grade Ss demonstrated that they could largely resolve the pronominal referent when real verbs were used. With the Bernoulli trials model, it was found that all sentences in the Each other type were ambiguous. For the other two sentence types, however, it was found that nearly all the sentences in the identical frames were unambiguous while nearly all the sentences in the nonsense frames were ambiguous. The set of unstarred cells (the sentences corresponding to these cells are interpreted as ambiguous) for the frames with real verbs contained sentences that were more complex syntactically. Hence the k-limited transducer model used this fact to explain the degeneration in the Ss' responses, i.e., the more syntactically complex sentences tended to exceed the memory span in the decoding stage of the two-stage model. With the nonsense He Xed him frame, the unstarred set contained 15 of the 16 cells; hence the k-limited transducer model explained the degeneration in which the memory span in the assignment stage was exceeded. Finally, the information transmission model computed the T-terms, which measured the dependency or predictability between the input sources and the response. The signigicant and the highly significant T-terms provided those combinations of the input sources which produced uniform responding.

In addition, since the T-terms were calculated as an average value over the 16 sentences of a frame, these served to summarize the Ss' responses for each sentence frame instead of individual sentences (of which there are 176).

5.3 The Seventh Grade Group

The basic data for the seventh grade group appear in Tables 5-15, 5-16 and 5-17. There were 13 male Ss with an average age of 12.5 years and an average I.Q. score of 115 and 13 female Ss with an average age also of 12.5 years and an average I.Q. score of 116. They were students at the West Junior High School in Ypsilanti, Michigan. Table 5-15 contains the results of three sentence frames of the Each other type; Table 5-16 contains the results of four sentence frames of the He Xed him back type; and Table 5-17 contains the results of four sentence frames of the He Xed him type. The results are tabulated separately for the male and female Ss and their pooled data are denoted as "Total". The difference between the male and female responses for each sentence frame is calculated as a chi-square value, which appears in the last column (under χ^2_{sex}) in Table 5-19. Since there was no significant difference between the male and the female responses for any of the sentence frames, subsequent discussion will pertain to the pooled data.

Each response matrix in Tables 5-15, 5-16 and 5-17 represents the 16 possible sentences in a sentence frame. The rows in each response matrix denote the grammatical forms of the first clause, while the columns denote the grammaticl forms of the second clause. Each cell of the response matrices corresponds to a particular sentence, and the number in the cell is the number of Ss who selected the grammatical object-slot as the referent of the pronoun he. For the nonsense frames, there were 12 female Ss instead of 13.

a. Analysis by the Bernoulli trials model. The maximum likelihood estimator $\hat{p}$ for each sentence frame (a matrix) appears at the bottom of the matrix. Next to the value of $\hat{p}$ is the confidence interval $[u,v]$ where the tail probability α is chosen as .01. Hence, if we assume a Bernoulli trial, i.e., that all the n Ss selected the grammatical object-slot as the referent of the pronoun he independently and with the same probability $\hat{p}$, then the probability of obtaining Y object-slots is $< .01$, if $Y \leq u$ or $Y \geq v$. Those cells whose numbers are $\leq u$ or $\geq v$ are starred. The sentences corresponding to these cells are interpreted as unambiguous sentences.

First, none of the cells in the response matrices in Table 5-15 are starred (except one, which is probably due to sampling error); hence all the sentences belonging to the Each other type are interpreted as ambiguous, again as we would expect. It is interesting to note that the estimated values of $\hat{p}$ (.56 to .60) are above .50, which shows a bias toward responding in the object-slot. The older groups (including the ninth grade group) were all biased toward responding in the subject-slot, and especially the college sophomore group, whose estimated values of $\hat{p}$ ranged from .26 to .28. It could be that when the structure of a sentence does not provide any cue for selecting the pronominal referent, adults tend to select the first mentioned person (i.e., a type of primacy effect) whereas younger children tend to select the person mentioned closest to the pronoun (i.e., a type of recency effect). However, the evidence presented so far is not convincing, even though the fifth grade's estimated $\hat{p}$ is in the range of .53 to .57. Perhaps this "recency effect" will appear more strongly when we examine the responses of much younger children, whose memory spans are exceeded by the sentence length.

Table 5-15

Response matrices for 13 male and 13 female[a] seventh grade subjects

Sentence type: John and Bill Yed each other and he Xed him.

Male

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	10	4	5	8
$\bar{A}$	5	8	5	6
$\bar{P}$	5	6	5	9
P	8	10	6	10

$\hat{p}$ = .53; [3,12]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	8	8	9	11
$\bar{A}$	5	6	7	3*
$\bar{P}$	7	7	9	9
P	10	11	4	6

$\hat{p}$ = .58; [3,12]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	7	7	8	8
$\bar{A}$	6	10	7	6
$\bar{P}$	7	9	9	9
P	5	7	7	9

$\hat{p}$ = .58; [3,12]

Female

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	7	6	10	7
$\bar{A}$	9	7	4	6
$\bar{P}$	7	6	7	8
P	8	8	5	9

$\hat{p}$ = .59; [3,12]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	9	7	9	10
$\bar{A}$	8	8	7	7
$\bar{P}$	9	9	9	6
P	10	7	8	8

$\hat{p}$ = .63; [4,13]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	5	6	10	6
$\bar{A}$	5	9	6	9
$\bar{P}$	9	9	7	6
P	6	10	8	6

$\hat{p}$ = .56; [3,12]

Total

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	17	10	15	15
$\bar{A}$	14	15	9	12
$\bar{P}$	12	12	12	17
P	16	18	11	19

$\hat{p}$ = .56; [8,21]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	17	15	18	21
$\bar{A}$	13	14	14	10*
$\bar{P}$	16	16	18	15
P	20	18	12	14

$\hat{p}$ = .60; [10,22]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	12	13	18	14
$\bar{A}$	11	19	13	15
$\bar{P}$	16	18	16	15
P	11	17	15	15

$\hat{p}$ = .57; [9,21]

a: Only 12 female Ss for the nonsense verbs

Starred cell (*) has $p < .01$

As for the responses from the sentence type He Xed him back, which appear in Table 5-16, a majority of the sentences were responded to randomly. We can discern a partition of the response matrix into corner submatrices only in the identical frame. Among the starred cells in the identical frame, the responses were clearly "diagonal", and, in addition, were consistent with the intended interpretation for the word "back". Without the added cues from pairs of identical verbs, however, this group of younger Ss was not able to assign the pronominal referent uniformly, except for the most simple active-active constructions. As for the logical reversal frame, we observe the "reversal" response, since the cells in the four corner submatrices have large numbers (e.g., greater than $n\hat{p}$) where the corresponding cells in the identical frame have low numbers (e.g., less than $n\hat{p}$), and vice versa. Of the 16 cells, only four of them attained the .01 level. Such a "reversal" response supports the hypothesis that the seventh grade children responded to the cues from the class of logical reversal verb-pairs, i.e., cues creating a previous action with a change of subject-object relationship. For the older Ss, a negative clause (either $\bar{A}$ or $\bar{P}$) followed by an affirmative clause (either A or P) tended to elicit a non-reversal responding. However, since the responses are largely random for the logical reversal frame (for a relatively small sample of 26 Ss), it is not clear whether the negative-affirmative constructions would also elicit non-reversal responding, especially in view of the fact that one of the four cells (cell $\bar{P}$-P) actually elicited a reversal response. We need to test again for this effect with a larger sample.

Finally we have the response matrices from the sentence type He Xed him, which appear in Table 5-17. Again a majority of the sentences were

Table 5-16

Response matrices for 13 male and 13 female[a] seventh grade subjects.
Sentence type: John Yed Bill and he Xed him back.

Male

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	11	11	7	7
$\bar{A}$	8	12*	6	5
$\bar{P}$	3*	6	6	6
P	7	6	3*	7

$\hat{p}$ = .53; [3,12]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	12*	11*	4	7
$\bar{A}$	8	9	2*	4
$\bar{P}$	6	3	7	5
P	6	6	6	6

$\hat{p}$ = .49; [2, 11]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	2*	2*	7	9
$\bar{A}$	5	5	9	6
$\bar{P}$	4	9	6	3
P	8	7	5	5

$\hat{p}$ = .44; [2,11]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	10	12*	5	4
$\bar{A}$	11	11	4	7
$\bar{P}$	6	8	9	11
P	5	2*	11	6

$\hat{p}$ = .59; [3,12]

Female

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	10	11	6	4
$\bar{A}$	6	9	6	5
$\bar{P}$	9	7	7	6
P	9	6	5	5

$\hat{p}$ = .58; [3,12]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	12*	10	6	5
$\bar{A}$	7	9	6	4
$\bar{P}$	7	5	7	8
P	4	5	6	9

$\hat{p}$ = .53; [3,12]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	2	1*	6	7
$\bar{A}$	4	2	7	6
$\bar{P}$	5	7	7	2
P	11*	5	5	5

$\hat{p}$ = .39; [1,10]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	13*	11	4	2*
$\bar{A}$	11	11	4	2*
$\bar{P}$	3*	8	7	6
P	11	2*	9	10

$\hat{p}$ = .55; [3,12]

Total

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	21*	22*	13	11
$\bar{A}$	14	21*	12	10
$\bar{P}$	12	13	13	12
P	16	12	8*	12

$\hat{p}$ = .56; [8,21]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	24*	21*	10	12
$\bar{A}$	15	18	8	8
$\bar{P}$	13	8	14	13
P	10	11	12	15

$\hat{p}$ = .51; [7,20]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	4*	3*	13	16
$\bar{A}$	9	7	16	12
$\bar{P}$	9	16	13	5*
P	19*	12	10	10

$\hat{p}$ = .42; [5,18]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	23*	23*	9*	6*
$\bar{A}$	22*	22*	8*	9*
$\bar{P}$	9*	16	16	17
P	16	4*	20	16

$\hat{p}$ = .57; [9,21]

a: Only 12 female Ss for the nonsense verbs
Starred cell (*) has $p < .01$

Table 5-17

Response matrices for 13 male and 13 female[a] seventh grade subjects.
Sentence type: John Yed Bill and he Xed him.

Male

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	9	7	3	5	10	10	6	8	2*	1*	5	11*	12*	12*	1*	2*
$\bar{A}$	5	7	6	9	9	10	5	4	9	3	11*	6	13*	12*	4	4
$\bar{P}$	6	7	8	5	8	8	7	8	6	8	3	5	5	6	9	10
P	8	4	6	8	8	5	7	8	10	10	5	4	4	6	11	10
	$\hat{p}$ = .50; [2,12]				$\hat{p}$ = .58; [3,12]				$\hat{p}$ = .48; [2,11]				$\hat{p}$ = .58; [3,12]			

Female

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	10	8	7	5	11	8	5	8	2*	2*	6	9	13*	13*	0*	3*
$\bar{A}$	8	10	5	4	7	9	3*	4	7	3	7	5	11	13*	5	1*
$\bar{P}$	5	9	8	7	8	3*	9	6	8	9	4	4	4	8	12*	10
P	7	7	8	8	7	5	8	9	8	6	7	5	7	2*	7	9
	$\hat{p}$ = .60; [3,12]				$\hat{p}$ = .53; [3,12]				$\hat{p}$ = .44; [2,11]				$\hat{p}$ = .57; [3,12]			

Total

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	19	15	10	10	21*	18	11	16	4*	3*	11	20*	25*	25*	1*	5*
$\bar{A}$	13	17	11	13	16	19	8*	8*	16	6*	18	11	24*	25*	9*	5*
$\bar{P}$	11	16	16	12	16	11	16	14	14	17	7	9	9*	14	21	20
P	15	11	14	16	15	10	15	17	18	16	12	9	11	8*	18	19
	$\hat{p}$ = .55; [8,20]				$\hat{p}$ = .56; [9,21]				$\hat{p}$ = .46; [6,19]				$\hat{p}$ = .58; [9,22]			

a: Only 12 female Ss for the nonsense verbs
Starred cell (*) has $p < .01$

responded to ambiguously. However, the identical frame produced a "diagonal" response which is consistent with responding to the deep structure of the sentences, and, in addition, follow the expected interpretation of "to interchange the two actors of the two actions". For the logical reversal frame, the responses are clearly "reversal", although there are only four starred cells. The remaining cells have small numbers when the corresponding cells in the identical frame have large numbers and vice versa. It is not clear how the Ss responded to the negative-affirmative constructions, since all these four deviant cells received random responding. For the nonsense frame, no cells are starred, indicating that the Ss were not able to gather enough cues from sentences with nonsense verb-pairs to properly assign the referent, not even from the most simple A-A sentence.

b. Analysis by the k-limited transducer model. In this section, we shall account for the observed degenerative behavior among the 16 sentences in each frame by the k-limited transducer model. Table 5-18 tabulates the 4-tuples (e,f,g,h) for each sentence frame in the He Xed him and the He Xed him back types.

We observe that only the identical frames produced maximal values in the 4-tuples, which means that the limits of the assignment transducer could not have been exceeded. However, as the sentences became more complex syntactically, the limits of the decoding transducer were exceeded, hence the drop from a maximum of four to two for element g and zero for element h of the 4-tuple (e,f,g,h). The 4-tuples for the similar and the logical reversal frames in both sentence types may be explained by a pair of k-limited transducers where the limits of the second (assignment) transducer were exceeded by some Ss, hence the non-maximal values for the first element of the 4-tuples; and where the limits of the first (decoding)

transducer were exceeded by passive constructions, hence the value zero or two for the other three elements of the 4-tuples. Finally, for the nonsense He Xed him frame, none of its 16 sentences was responded to uniformly, that is, there were not enough cues to produce a correct

Table 5-18

Number of starred cells in each corner submatrix for the seventh grade subjects

	Nonsense	Similar	Log. Rev.	Identical
He Xed him	0,0,0,0	1,2,0,0	3,1,0,0	4,4,2,0
He Xed him back	3,0,0,1	2,0,0,0	2,0,1,1	4,4,2,0

assignment. The all-random responses may be accounted for by a pair of k-limited transducers where the limits of the second (assignment) transducer were exceeded by all the Ss. For the nonsense He Xed him back frame, the cues from the word "back" made the assignment much easier, at least for the simple active-active constructions. But when sentences contained passive grammatical forms the assignment became random, probably due to incorrect decoding of the sentences into their respective deep structures. Hence we can explain the degeneration of this nonsense He Xed him back frame with the k-limited transducer model where the limits of the second transducer were not exceeded, but where the limits of the first (decoding) transducer were exceeded whenever there was a passive transformation.

c. Analysis by the information transmission model. The results of the analysis by information transmission model appear in Table 5-19. The columns are components of transmitted information (the T-terms) and interaction (the A-terms) terms. Again any checked (√) term is highly sign-

Table 5-19

Components of transmitted information for 13 male and 13 female[a] seventh grade subjects. The average I.Q. scores for the male and female Ss are 115 and 116 respectively.

		H(Z)	T(ABC)	T(AB)	T(AB)/M	T(AB)/F	T(A)	T(B)	T(C)	A(ABZ)	A(ACZ)	A(BCZ)	A(ABCZ)	χ^2_{sex}
	DF		31	15	15	15	3	3	1					16
He Xed him	Nonsense	.99	6.57	3.33	5.07	6.39	.02	.42	.87	2.90	.14	.62	1.61	8.81
	Similar	.99	7.65	6.16*	5.64	9.25	.79	1.19	.21	4.18	.17	.27	.85	4.03
	Log. Rev.	1.00	14.00✓	12.16✓	18.77✓	9.06	1.08	.36	.08	10.72	.37	.03	1.36	4.47
	Identical	.98	36.00✓	32.09✓	30.57✓	41.39✓	.39	3.27✓	.02	28.43	.44	.06	3.40	6.77
He Xed him back	Nonsense	.99	10.54*	8.03✓	11.82*	8.84	1.65	3.07✓	.15	3.31	.80	.49	1.07	7.02
	Similar	1.00	10.84✓	9.24✓	12.64*	8.81	1.77*	1.42	.11	6.05	.20	.28	1.01	4.74
	Log. Rev.	.98	10.94✓	9.62✓	9.41	12.12*	.79	.76	.18	8.07	.21	.45	.48	2.98
	Identical	.99	22.96✓	18.26✓	18.00✓	27.70✓	.12	2.18*	.11	15.96	1.20	.78	2.61	14.98
Each other	Nonsense	.98	5.91	3.23	6.67	4.87	.49	.44	.10	2.29	.18	.14	2.26	8.16
	Similar	.97	6.29	3.76	9.70	2.44	1.52	.14	.22	2.11	.39	.20	1.72	7.09
	Log. Rev.	.99	4.29	2.51	3.17	5.36	.29	1.06	.03	1.17	.13	.15	1.47	3.57

a: Only 12 female Ss for the nonsense frames
Checked item (✓) has p < .001
Starred item (*) has p < .01

ificant, i.e., under the null hypothesis of no transmission between the input sources and the output source, the probability of obtaining the term of such a magnitude is $< .001$. Similarly, any starred (*) term is significant when the probability of obtaining such a magnitude under the null hypothesis is $< .01$.

The T-terms for the three frames in the <u>Each other</u> type are all nonsignificant. Thus the output response was independent from the input sources, i.e., any knowledge of the gramatical form of the first clause (input source A), the grammatical form of the second clause (input source B) and/or the sex of the <u>S</u>s (input source C) did not influence the prediction of the pronominal referent (output source Z). This is consistent with the finding from the Bernoulli trials model that each of the 16 sentences in any of the <u>Each other</u> frames was a random response.

for the four frames of the <u>He Xed him back</u> type, the T(AB) terms are all highly significant, while some of the other doubly-controlled and triply-controlled T-terms are either significant or nonsignificant. Thus the group as a whole was able to use the cues from the word "back" for uniform responding for any class of verb-pairs. However, when the male and female responses are analyzed separately, although the T(AB)/M or T(AB)/F terms are of the same magnitude as T(AB), some of these are not significant. Since the χ^2 value for the terms is directly proportional to the sample size n ($\chi^2 = 1.3863$ n T), a reduction in n gives a corresponding reduction in the χ^2 value. Some of the T(A) and T(B) terms are also significant or highly significant, which suggest some bias in the resulting responses when summed over the other two input sources. The maximum T(AB) term comes from the identical frame. Its value of 18% corresponds to an m of 7 in Table C-1. Thus we can expect at most a $(26-7)/26 = 73\%$ agreement among the 26 seventh grade <u>S</u>s in

responding to the <u>He Xed him back</u> type.

As for the <u>He Xed him</u> type, only the identical and the logical reversal frames have highly significant T-terms. Hence in the absence of the cues from the word "back", a knowledge of the grammatical forms of the two clauses in the nonsense frame could not predict the response, and the prediction is significant (though not highly significant) for the similar frame. The maximum T(AB) term comes from the identical frame. Its value of 37% corresponds to an m of 5 in Table C-1. Thus we can expect at most a (26-5)/26 = 81% agreement among the 26 seventh grade <u>S</u>s in responding to the <u>He Xed him</u> type.

Again we observe a larger maximum agreement when the word "back" was absent in sentences with identical verb-pairs. Thus the double cues from "back" and "identical verbs" produced less uniform response than just "identical verbs". On the other hand, the cues from "back" alone produced much more uniform responding with either the nonsense verbs or the similar verbs than without the word "back'. The maximum of 81% agreement for the identical <u>He Xed him</u> frame is actually higher than the corresponding one among the eighth grade <u>S</u>s, which is 78%. The relatively small difference between 81% and 78% produced a larger difference in the maximum T(AB;Z) of 32% and 23% may nevertheless be a sampling error.

Again, we wish to know whether the <u>S</u>s based their responses on the sentence type or on the particular verb-pairs. Hence we want to find out whether there is any difference between responses from two classes of verb-pairs for the same sentence type and whether there is any difference between the responses from any two sentence types with the same class of verb-pairs. In Table 5-20, we have tabulated the $\chi^2(16)$ values between any two sentence types for each class of verb-pairs. As with the eighth grade

Table 5-20

Chi-square values between any two sentence types for each class of verb-pairs for the seventh grade subjects.

	Xed him vs Xed him back	Xed him vs Each other	Xed him back vs Each other
Nonsense verbs	9.0	9.7	22.1
Similar verbs	4.3	16.4	32.6*
Log. Rev. verbs	9.0	43.0√	40.4√
Identical verbs	13.3		

Checked item (√) has $p < .001$
Starred item (*) has $p < .01$

Table 5-21

Chi-square values between any two verb-pairs in each sentence type for the seventh grade subjects.

	NS-SI	NS-LR	NS-ID	SI-LR	SI-ID	LR-ID
He Xed him	6.7	55.6√	45.6√	68.6√	40.6√	164.8√
He Xed him back	7.6	70.3√	23.3	80.4√	21.3	113.8√
Each other	6.8	8.8		13.0		

Checked item (√) has $p < .001$

Ss, there was no significant difference between the He Xed him and the He Xed him back types. However, in contrast with the two older groups, there is no significant difference among the three sentence types with nonsense verb-pairs, as well as with the similar frames of the He Xed him and the Each other types. Since the responses to the Each other type were interpreted as random, the responses from the nonsense He Xed him, the nonsense He Xed him back, and the similar He Xed him frames were also essentially random, as these were not significantly different from the Each other type.

Table 5-21 tabulates the $\chi^2(16)$ values between any two verb-pairs in each sentence type. Again we note that the Ss responded to the Each other sentence type in a way that was independent of the particular class of verb-pairs, since all the sentences elicited random responding. Large significant differences still come from NS-LR, SI-LR and LR-ID pairs, which confirm our finding from the Bernoulli trials model that the response matrices using the logical reversal verb-pairs in the He Xed him and He Xed him back types included the "reversal" response, although both had only four starred cells. The only other highly significant terms are the NS-ID and SI-ID pairs in the He Xed him type. Thus in the absence of the word "back", the identical frame produced a unique response which is different from the nonsense and similar (and, or course, also the logical reversal) frames. However, when the word "back" was present, all the nonsense, similar and identical frames in the He Xed him back type were responded to alike.

d. Summary. The results from the seventh grade Ss showed a further degeneration with respect to the responses of the eighth grade Ss. These

may be attributed to shorter memory spans, both in the decoding and in the assignment stages of the k-limited transducer model. With the Bernoulli trials model, it was found that all the cells in the Each other type were not starred, and the sentences corresponding to these were interpreted as ambiguous. For the other two sentence types, only the identical frames have more than half of the cells starred, and the sentences corresponding to these starred cells were interpreted as unambiguous. Although only four cells in the logical reversal frames were starred, the direction in which these cells attained the .01 significance level, i.e., whether the cell number was $\leq u$ or $\geq v$, indicated that the response matrix included "reversal" responses. The k-limited transducer model also used the assumed shorter memory span of seventh grade Ss to explain the degeneration of response across the 16 sentences of a frame. Finally with the information transmission model, it was found that the T(AB) terms for all four frames in the He Xed him back type were highly significant while only the identical the the logical reversal frames in the He Xed him frame were highly significant. That is to say, we can predict the output response when the grammatical form of the two clauses are known in these six frames. In addition the maximum percentage agreement for the identical He Xed him frame turned out to be higher for seventh than for eighth grade Ss.

5.4 The Fifth Grade Group

The basic data for the fifth grade group appear in Table 5-22, 5-23 and 5-24. There were 14 male Ss with an average age of 10.7 years and an average I.Q. score of 98, and 9 female Ss with an average age

of 10.8 years and an average I.Q. score of 105. There were two Ss, one male S and one female S, who were not able to read. The teacher read the test material, one sentence at a time, to the male non-reader and helped him mark an "x" on top of the word he orally selected, while the writer did the same with the female non-reader. When their responses were tabulated, there were no systematic differences with the other Ss in this age group, and their responses were pooled along with the rest of the fifth graders. The Ss were students at the Adams Elementary School in Ypsilanti, Michigan.

Each response matrix in Tables 5-22, 5-23 and 5-24 represents the 16 possible sentences in a sentence frame. The rows in each matrix denote the grammatical forms of the first clause, and the columns denote the grammatical forms of the second clause. Each cell of the response matrices corresponds to a particular sentence, and the number in the cell is the number of Ss who selected the gramatical object-slot as the referent of the pronoun he. The male and the female responses are tabulated separately with their pooled responses as "Total". Since the male and the female responses are not significantly different, subsequent discussion will pertain to only the "Total" matrices.

a. Analysis by the Bernoulli trials model. The maximum likelihood estimator $\hat{p}$ for each sentence frame (a matrix) appears at the bottom of the matrix. Next to the value of $\hat{p}$ is the confidence interval [u,v] where the tail probability α is chosen as .01. Hence, if we assume a Bernoulli trials model, i.e., all the n Ss selected the grammatical object-slot as the referent of the pronoun he independently and with the

Table 5-22

Response matrices for 14 male and 9 female fifth grade subjects

Sentence type: John and Bill Yed each other and he Xed him.

Male

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	8	10	7	8
$\bar{A}$	9	10	11	8
$\bar{P}$	9	11	8	8
P	7	9	7	8

$\hat{p}$ = .62; [4,14]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	9	5	10	6
$\bar{A}$	10	4	7	6
$\bar{P}$	9	9	7	8
P	7	8	9	5

$\hat{p}$ = .53; [3,13]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	6	5	10	12
$\bar{A}$	9	7	7	7
$\bar{P}$	8	6	6	8
P	7	7	7	9

$\hat{p}$ = .54; [3,13]

Female

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	6	7	2	5
$\bar{A}$	4	6	3	3
$\bar{P}$	5	5	4	5
P	3	7	5	3

$\hat{p}$ = .51; [1,9]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	5	5	7	4
$\bar{A}$	4	6	5	4
$\bar{P}$	3	7	2	4
P	4	5	5	5

$\hat{p}$ = .52; [1,9]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	6	7	3	6
$\bar{A}$	4	4	7	5
$\bar{P}$	5	4	6	4
P	3	6	5	6

$\hat{p}$ = .56; [2,9]

Total

Nonsense Verbs

	A	A	P	P
A	14	17	9	13
$\bar{A}$	13	16	14	11
$\bar{P}$	14	16	12	13
P	10	16	12	11

$\hat{p}$ = .57; [8,19]

Similar Verbs

	A	A	P	P
A	14	10	17	10
$\bar{A}$	14	10	12	10
$\bar{P}$	12	16	9	12
P	11	13	14	10

$\hat{p}$ = .53; [7,19]

Logical Reversal Verbs

	A	A	P	P
A	12	11	13	18
$\bar{A}$	13	11	14	12
$\bar{P}$	13	10	12	12
P	10	13	12	15

$\hat{p}$ = .55; [8,19]

Table 5-23

Response matrices for 14 male and 9 female fifth grade subjects

Sentence type: John Yed Bill and he Xed him back.

Male

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	10	8	7	10	9	9	10	5	7	9	10	11	9	10	11	8
$\bar{A}$	7	8	6	9	13*	9	9	10	7	9	7	7	11	14*	10	6
$\bar{P}$	9	7	9	7	8	7	7	7	9	8	4*	8	7	7	5	8
P	9	6	4	8	9	8	9	6	9	9	8	12	6	4*	8	9
	$\hat{p}$ = .55; [3,13]				$\hat{p}$ = .60; [4,13]				$\hat{p}$ = .60; [4,13]				$\hat{p}$ = .59; [4,13]			

Female

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	5	6	4	4	5	4	5	5	4	4	5	5	8	6	4	5
$\bar{A}$	6	6	5	4	2	5	6	4	3	3	2	4	5	6	7	4
$\bar{P}$	4	5	6	4	4	4	3	4	5	2	5	4	3	5	4	6
P	3	5	4	2	3	5	3	2	7	7	4	6	6	3	5	5
	$\hat{p}$ = .51; [1,9]				$\hat{p}$ = .44; [1,8]				$\hat{p}$ = .49; [1,9]				$\hat{p}$ = .57; [2,9]			

Total

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	15	14	11	14	14	13	15	10	11	13	15	16	17	16	15	13
$\bar{A}$	13	14	11	13	15	14	15	14	10	12	9	11	16	20*	17	10
$\bar{P}$	13	12	15	11	12	11	10	11	14	10	9	12	10	12	9	14
P	12	11	8	10	12	13	12	8	16	16	12	18	12	7*	13	14
	$\hat{p}$ = .53; [7,19]				$\hat{p}$ = .54; [7,19]				$\hat{p}$ = .56; [7,19]				$\hat{p}$ = .58; [8,20]			

Starred cell (*) has $p < .01$

Table 5-24

Response matrices for 14 male and 9 female fifth grade subjects

Sentence type: John Yed Bill and he Xed him.

Male

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	11	9	8	10	9	6	6	7	7	9	12*	10	10	8	8	9
$\bar{A}$	5	10	8	7	9	7	8	6	4	8	8	7	9	12	9	4*
$\bar{P}$	7	10	8	7	8	8	9	10	3*	7	8	7	7	8	8	10
P	6	3*	8	6	9	10	9	7	7	8	5	5	5	7	12	5
	$\hat{p}$ = .55; [3,13]				$\hat{p}$ = .57; [4,13]				$\hat{p}$ = .51; [3,12]				$\hat{p}$ = .58; [4,13]			

Female

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	3	4	4	6	4	5	5	4	1*	2	4	6	7	8	4	3
$\bar{A}$	6	4	5	4	6	3	5	3	3	3	4	5	7	5	4	3
$\bar{P}$	4	2	8	3	1*	5	4	4	3	5	7	3	4	4	2	5
P	6	6	2	7	4	4	7	4	6	3	6	2	3	3	4	5
	$\hat{p}$ = .51; [1,9]				$\hat{p}$ = .47; [1,9]				$\hat{p}$ = .44; [1,8]				$\hat{p}$ = .44; [1,9]			

Total

	Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
A	14	13	12	16	13	11	11	11	8	11	16	16	17	16	12	12
$\bar{A}$	11	14	13	11	15	10	13	9	7	11	12	12	16	17	13	7*
$\bar{P}$	11	12	16	10	9	13	13	14	6*	12	15	10	11	12	10	15
P	12	9	10	13	13	14	16	11	13	11	11	7	8	10	16	10
	$\hat{p}$ = .54; [7,19]				$\hat{p}$ = .54; [7,19]				$\hat{p}$ = .49; [6,18]				$\hat{p}$ = .55; [7,19]			

Starred cell (*) has $p < .01$

same probability $\hat{p}$, then the probability of obtaining Y object-slots is $< .01$, if $Y \leqslant u$ or $Y \geqslant v$. Those cells whose numbers are $\leqslant u$ or $\geqslant v$ are starred. The sentences corresponding to these cells are interpreted as unambiguous sentences.

Of the 176 cells in the 11 response matrices, only four cells are starred, i.e., there were four unambiguously interpreted sentences by the fifth grade subjects. As expected, all the sentences from the Each other sentence type were ambiguous. As for the He Xed him back type, only two sentences from the identical frame were unambiguous, while for the He Xed him type, only one sentence each from the identical and the logical reversal frames were unambiguous. Thus, for all practical purposes, the fifth grade Ss were not able to resolve the ambiguity of the pronominal referent; that is, they were not able to use the cues from either the word "back" or the verb-pairs to produce a uniform response.

b. Analysis by the k-limited transducer model. Again this model is intended to account for the observed degenerative behavior among the 16 sentences in each frame. Table 5-25 tabulates the 4-tuples (e,f,g,h) for each sentence frame in the He Xed him and the He Xed him back types. Since this is our youngest age group, we can assume that the Ss have shorter memory spans, on the average, than the other groups, Assuming shorter memory spans, we may account for the nearly complete random responding in the following way. Since the A-A sentence contains clausses whose surface structure is the same as its deep structure, its demand on decoding is minimal. Fraser, et al [1963] found that children of age three were able to decode active sentences; hence it seems reasonable to expect children of age 10.7 (average age of the fifth graders) would be

able to decode the A-A sentences. Yet the Ss were not able to respond uniformly to this simplest construction; therefore it seems safe to assume that the difficulty was not in the decoding stage, but rather in the

Table 5-25

Number of starred cells in each corner submatrix for the fifth grade subjects

	Nonsence	Similar	Log. Rev.	Identical
He Xed him	0,0,0,0	0,0,0,0	0,0,1,0	0,1,0,0
He Xed him back	0,0,0,0	0,0,0,0	0,0,0,0	1,0,1,0

assignment. We can then account for the observed data by the k-limited transducer model where the limits of the k' in the assignment stage were exceeded.

c. Analysis by the information transmission model. The results from the analysis by information transmission model appear in Table 5-26. The columns are components of transmitted information (the T-terms) and interaction (the A-terms) terms. Again any starred (*) term is significant i.e., under the null hypothesis of no transmission between the input sources and the output source, the probability of obtaining the term of such a magnitude is $< .01$.

The only doubly-controlled T-terms that are significant come from the identical He Xed him back frame. Hence under the optimum condition of having cues both from the word "back" and from "identical verbs", this youngest group of Ss was able to produce a significantly uniform response. The T(A) term for this frame is also significant, which suggests that knowing either the grammatical form of the first clause or knowing both

Table 5-26

Components of transmitted information for 14 male and 9 female fifth grade subjects. The average I.Q. scores for the male and female Ss are 98 and 105 respectively.

		H(Z)	T(ABC)	T(AB)	T(AB)/M	T(AB)/F	T(A)	T(B)	T(C)	A(ABZ)	A(ACZ)	A(BCZ)	A(ABCZ)	χ^2_{sex}
	DF		31	15	15	15	3	3	1					16
He Xed him	Nonsense	1.00	8.20	2.15	6.32	10.91	.52	.06	.09	1.56	1.36	.24	4.37	16.71
	Similar	1.00	4.85	2.19	2.66	6.52	.33	.29	.68	1.57	.53	.44	1.02	6.02
	Log. Rev.	1.00	9.02	4.90	7.27	10.71	.49	1.73	.40	2.68	1.51	.47	1.74	10.01
	Identical	.99	9.32	5.52	7.87	10.07	.84	.56	.59	4.12	.22	.87	2.12	8.85
He Xed him back	Nonsense	1.00	4.20	1.92	3.71	4.59	.83	.32	.15	.78	.24	1.18	.71	4.37
	Similar	1.00	6.92	2.22	5.49	4.68	1.12	.54	1.74*	.57	.52	.61	1.83	13.35
	Log. Rev.	.99	7.16	4.18	5.21	7.94	2.23	.63	.88	1.32	.40	.23	1.46	5.58
	Identical	.98	9.54	6.51*	11.21*	6.83	2.48*	.10	.04	3.93	.33	.10	2.56	6.65
Each other	Nonsense	.98	5.58	2.89	2.63	7.99	.18	1.90	.85	.81	.45	.29	1.10	6.06
	Similar	1.00	5.32	2.93	4.90	5.94	.11	.52	.01	2.30	.43	1.17	.78	5.50
	Log. Rev.	.99	4.78	2.09	4.55	5.09	.21	.69	.02	1.19	.01	.32	2.35	5.96

Starred item (*) has $p < .01$

Table 5-27

Chi-square values between any two sentence types for each class of verb-pairs for the fifth grade subjects.

	Xed him vs Xed him back	Xed him vs Each other	Xed him back vs Each other
Nonsense verbs	1.2	7.7	4.4
Similar verbs	5.6	4.2	4.4
Log. Rev. verbs	18.5	12.3	5.7
Identical verbs	4.9		

Table 5-28

Chi-square values between any two verb-pairs in each sentence type for the fifth grade subjects.

	NS-SI	NS-LR	NS-ID	SI-LR	SI-ID	LR-ID
He Xed him	8.2	8.3	11.6	12.9	8.4	25.6
He Xed him back	5.1	12.1	12.9	15.3	9.8	23.1
Each other	11.3	10.0		9.8		

clauses can predict the response better than chance. The generally non-significant results from the information transmission model imply that the Ss did not respond according to the structure of the sentence (except those in the identical He Xed him back frame). This finding is consistent with the finding from the Bernoulli trials model that only four sentences out of a total of 176 were interpreted as unambiguous.

In Table 5-27 and 5-28, the $\chi^2(16)$ values are tabulated for the differences between any two sentence types for each class of verb-pairs and between any two classes of verb-pairs in each sentence type. Since there were only four starred cells in the Bernoulli trials model and the output response was found to be independent of the input sources (except for the identical He Xed him back frame) in the information transmission model, it should not be surprising that all the pair-wise differences are small and nonsignificant. In fact even the logical reversal verb-pairs were responded to no different from the other verb-pairs.

d. Summary. In general, the fifth grade Ss were not able to resolve the pronominal ambiguities. For this youngest group, the nearly complete random responding was characterized by the k-limited transducer model as a case in which the limits of memory span were exceeded. The Bernoulli trials model found that nearly all the sentences in every frame were random, while the information transmission model found that the output responses were independent from the input sources. All models support the conclusion that the fifth-graders could not resolve the pronominal ambiguity.

Chapter 6

DISCUSSION AND SUMMARY

We have proposed three mathematical models to analyze the observed data in the form of contingency tables. The Bernoulli trials model takes the responses to a block (in this case, a frame of 16 sentences) of sentences and applies probability theory to determine the probability level of each response. Three assumptions must be satisfied before we can apply this model:

1) Each trial has only two outcomes.

2) Each trial is independent from any other trial.

3) The probabilities of the two outcomes remain constant in all trials.

We are treating each trial as "one S responding to one sentence". Since the Ss were instructed to mark an "x" on top of either John or Bill (and not both) for every sentence, each trial has exactly two outcomes. Since each S was working on his own questionnaire booklet, it is reasonable to assume that one S's responses were independent from any other S's responses. The third assumption requires that each S in a group has the same probability in selecting, for example, the word John. It is impossible to know whether this assumption is, or can ever be, satisfied. The procedure used was to apply the best available estimator, the maximum likelihood estimator $\hat{p}$, to a large sample. The sample selected was the 16 different sentences in a frame, giving a total of 16n sample points where n is the number of Ss in a group.

This estimated $\hat{p}$ does not distinguish between a group of Ss in which all of them truly were neutral in their selection of John or Bill as the referent of the pronoun he in a sentence and a group where half of the Ss strongly felt the selection was Bill and the other half strongly felt the selection was John. Assuming that the Ss in both groups were consistent in treating the 16 sentences in a frame, then both groups would give a $\hat{p}$ close to .50. Fortunately, we need not distinguish the two cases, since we are concerned with whether the sentences are ambiguous or not.[3] In both groups, since the sentences had elicited non-uniform responses, the sentences must be considered ambiguous. However, the estimated $\hat{p}$ along with the confidence interval $[u,v]$ can distinguish the above group responses from one in which the subject-slot was uniformly chosen in eight of the 16 sentences and the object-slot was chosen in the remaining eight sentences (e.g., the "diagonal" response). Again the estimated $\hat{p}$ is close to .50, but since each of the individual cell numbers is either high (uniform object-slot) or low (uniform subject-slot), they will all fall outside the confidence interval, i.e., either $\leqslant u$ or $\geqslant v$. Since the selected tail-probability α is .01, then the probability for the number in a cell being $\leqslant u$ or $\geqslant v$ is $< .01$, i.e., it is not likely that the response to that cell was the result of Bernoulli

---

[3] For those who are interested in the Ss' response behavior in an all-or-none learning situation, the distinction is important. Under such circumstances, a more complex experiment could be designed in which each S is asked to give his response together with a three-point or five-point scale for his subjective strength for the response. Without such a subjective strength scale, it is impossible to distinguish with a matrix of zeros and ones (response of one S to one sentence frame) between the result of random selection and the result of language usage. Any matrix of zeros and ones is equi-probable to any other matrix of zeros and ones. But when a group of such matrices are summed, the numbers in each cell follow a binomial distribution, if each matrix was the result of random selection from each S. If the cell numbers deviate significantly from a binomial distribution, as measured by the confidence interval $[u,v]$, then we have some confidence that the responses were based on language usage, and therefore, the sentences are not ambiguous.

trials. Hence we interpret the sentence that produced such a response as an unambiguous sentence.

The k-limited transducer is a precise mathematical model; however, when applied to the present data, it serves as conceptual means to explain degenerative responding, either when sentences become more complex syntactically, or when Ss become younger. Since we do not know the psychological unit, whether it be the morpheme, the word, or some higher structured linguistic unit, that the Ss used to analyze each sentence and to select the pronominal referent, we cannot be very precise about applying the model. The basic assumptions of the model are that the internal states of the transducer are finite and that each transition depends on reading the input tape of at most k units long. These two assumptions are easily satisfied, since the human brain is definitely finite and there is surely a certain maximum length of sentence beyond which Ss cannot process correctly.

In their experiment on the recall of sentences with various degrees of self-embedding clauses, Miller and Isard [1964] postulated a type of push-down memory for explaining their findings. Similarly, McNeill [1965] postulated three separate memory spans in explaining the findings of Fraser, et al [1963] and Ervin [1964]. For the present data, the writer postulated two separate memory spans in the k-limited transducer model. The model, as diagramed in Figure 4-1, uses two k-limited transducers in sequence: the first one decodes each sentence into its deep structure and the second one applies the S's knowledge of English usage to assign a pronominal referent. Even though the present experiment was conducted in written form, memory limitations could still contribute to degenerative responding. The written form does not require the S to store each sentence in his immediate memory for later processing. However, in order to select the most appropriate referent for the pronoun, the S must analyze each sentence to

obtain the subject-object relationship in the two clauses. The results of intermediate analysis must be stored so that the pronominal referent could be selected in accordance with conventional usage.

The sequential arrangement of a decoding transducer, followed by an assignment transducer, was strongly influenced by the procedure generally accepted in the area of language processing by computer [Simmons, 1965; Hays, 1963; Lamb, 1965b], of a syntactic-analysis program followed by a semantic-analysis program. This arrangement was also suggested by the theoretical paper by Katz and Fodor [1963], in which semantic rules are applied only after a complete syntactic analysis to obtain the meaning of a sentence. The decoder transducer corresponds to a syntactic analysis in the sense that only grammatical rules are applied to obtain the deep structure of a sentence while the assignment transducer corresponds (at least in part) to a semantic analysis in the sense that the knowledge of the English language is applied to select the most appropriate pronominal referent.

The k and the k' in the two transducers denote specific length of input strings. But sometimes we observe different responses when two sentences are of the same length, either by word or morpheme count, as for example in <u>John bejed Bill and he was gowed by him</u> vs. <u>John punched Bill and he was punched by him</u>. Hence we hypothesize a decoding "program" that is concatenated with the sentence as an input string to the decoding transducer and an assignment "program" that is concatenated with the decoded string (from the decoding transducer) as an input string to the assignment transducer. The "program" reflects a level of difficulty that is caused by the level of syntactic complexity in a sentence and the level of association strength between the two verbs of a particular verb-pair. As mentioned in Chapter 4, there is some behavioral evidence from the re-

sults of Mehler [1963], Mehler and Miller [1964], and Savin and Perchonock [1965] for including a decoding "program", and from Epstein [1961] and Rosenberg [1966] for including an assignment "program".

The weakest part of this model is perhaps the serial ordering of the two transducers. The writer does not know of any behavioral evidence in support of the hypothetical ordering, that of syntactic analysis followed by semantic analysis, nor, indeed, does he believe that it ever could be shown to be true. A recent experiment by Hull and Gough [1966] attempted to test the hypothesis that a listener analyzes a sentence syntactically before he analyzes it semantically. But due to interactions between syntactic and semantic analyses, the authors concluded that their data did not support the hypothesis. While the ordering of the two analyses facilitate both computer processing and theory construction, the human being is not necessarily constrained in this manner.

The development of the information transmission model as presented in Chapter 4 is directly applicable to the data in contingency tables. The advantage of this model is that each of the design variables is treated as an independent input source from which the relation between the individual input sources or any of these in combination, and the output source can be computed. Such a relation, measured by the T-terms (transmitted information terms), gives us the degree of predictability of the output response, i.e., the uniformity of responding in a group.

We did not use all the design variables as input sources for the information transmission model, because the model computes an average value over the different values of a variable. For example, if the sentence types were included as in input source, then (along with the other variables) the transmitted information terms could be computed. However, since the responses to the Each other type were generally random and the responses to

the <u>He Xed him</u> and to the <u>He Xed him back</u> types were generally "diagonal", then the average of these three sentence types would not correctly represent any sentence.

The results of the present experiment demonstrated that <u>S</u>s could reduce the number of interpretations of a key word in a sentence. In this we agree with the results of Kaplan's experiment [1949]. But unlike Kaplan's results, the present experiment showed that the resolution of the pronominal referent was not contingent on neighboring words. The parameters that could explain the different response matrices for the pronominal referent were the subject-object relationship and the two possible interpretations: "the same person performed both actions" and "each person performed one action". The design variables of four grammatical forms, three sentence types and four classes of verb-pairs caused differences in response matrices due to a change in either the subject-object relationship or one of the two possible interpretations. The changes caused by the grammatical forms were obvious, since a passive construction alternated the subject-object relationship from the corresponding active construction, while a negative construction generally did not, since the subject-object relationship was invariant under the negative transformation. The sentence types could make one of two interpretations more probable, or neither one as in the <u>Each other</u> type. The changes caused by the different classes of verb-pairs might be attributed to a class of semantic features to indicate a direction of action for all transitive verbs. At least three of these should be considered. There is an <u>outward</u> feature, exemplified by "hit", "strike", and a majority of transitive verbs, where the action is from the subject to the object; an <u>inward</u> feature, exemplified by the first word of the logical reversal verb-pairs, "understand", "remember", etc., where the subject receives (or senses) an action; and a <u>reciprocal</u> feature, exemplified by "meet", "marry",

etc., where the action between the subject and the object is reciprocated, i.e., A Xed B implies B Xed A. Thus for a given interpretation, the response in which the first verb of the verb-pair has the "inward" feature is expected to be opposite from one in which the first verb of the verb-pair has the "outward" feature, and in addition, the response in which the first verb of the verb-pair has the "reciprocal" feature is expected to be random. As a consequence, the dictionary for computer processing, using the English language, should contain such a class of semantic feature for the transitive verbs.

There are parameters which caused degenerative (i.e., tendency toward random responding) behavior. Any sentence from the Each other type evoked a random response. A syntactically complex sentence (involving the passive transformation) evoked less uniform responding than more simply constructed sentences. Changing from identical verb-pairs to the other classes of verb-pairs caused degeneration in Ss'responding, and a similar degeneration was observed when real verb-pairs were changed to nonsense verb-pairs. In general, deletion of the word "back" also caused some degeneration, especially with nonsense verb-pairs. Finally, the younger the Ss, the more degenerative their responses; at the extreme, the youngest group, wigh an average age of 10.7 years, produced random responding to nearly all the sentences. All these degenerative responses can be explained with the k-limited transducer model. If a sentence exceeds either the k units available for decoding, or the k' units available for assigning the pronominal referent, or both of these, then that sentence evokes a random response.

The results revealed that the fifth grade Ss could not resolve pronominal ambiguity. Further experimentation with this age group may isolate the point of difficulty. Then by varying the verb-pair, simplifying the construction of the sentence and/or providing some situational cues, such

as a short story of which the sentence is a part, it might be possible to detect the syntactic and semantic features needed in the process of resolving pronominal ambiguity. Presumably these features are learned as Ss are more exposed to everyday language usage. Likewise these features, in addition to others, must necessarily be included in a computer program, if it is to simulate human language processing.

Two developmental changes occurred sometime between the fifth and eighth grades: one in acquiring the semantic feature for "inward" and "outward" verbs, and the other in acquiring the constructional ambiguity when sentences using the "inward" verbs are expressed in the negative-affirmative constructions (i.e., sentences corresponding to the "deviant" cells in the logical reversal frames). Since the responses from the fifth grade Ss were all random, it is safe to assume that these youngest Ss could not distinguish the semantic feature for the "inward" and the "outward" verbs. On the other hand, the seventh grade Ss could make such a distinction, but the results were not conclusive on whether these Ss also acquired the constructional ambiguity. Thus with individual testing in this age range, where the experimenter could give additional cues, it might be possible to isolate the features that the Ss must necessarily learn before the distinctions were made. Again, for computer applications, such features must somehow be encoded in the program before it could claim to have any "intelligence".

Finally, the results revealed that the sentences with nonsense verb-pairs were responded to as if the verbs had the "outward" semantic feature. This is perhaps not surprising in view of the fact that most transitive verbs have this feature. However, it is an interesting empirical question whether this would still be true of the nonsense verbs if sentences con-

taining them were presented immediately following a set of sentences whose first verbs all have the "inward" feature, viz., the set of logical reversal verbs, or following a set of sentences whose first verbs all have the "reciprocal" feature. For sentences with identical verb-pairs, responses did not depend on whether or not the sentence included the word "back", but for sentences with nonsense verb-pairs, the responses degenerated considerably when the word "back" was absent. If we were to test sentences when the two nonsensical verbs were identical, the present finding would predict no degeneration when the word "back" is absent.

BIBLIOGRAPHY

Atkinson, R. C. and Shiffrin, R. M. [1965] Mathematical models for memory and learning. Technical Report Number 79, Institute for Mathematical Studies in the Social Science, Stanford University.

Bar-Hillel, Y. [1964] Language and Information: Selected Essays on Their Theory and Application. Reading, Mass.: Addison-Wesley.

Bobrow, D. G. [1963] Syntactic analysis of English by computer - a survey. Proceedings Fall Joint Computer Conference, 24, 365-388. Baltimore: Spartan Books.

Bollinger, D. [1965] The atomization of meaning. Language, 41, 555-573.

Brown, R. and Fraser, C. [1964] The acquisition of syntax, in The Acquisition of Language, Monographs of the Society for Research in Child Development, Volume 29, Number 1, Serial Number 92, 43-79. Ursula Bellugi and R. Brown (Editors). Yellow Springs, Ohio: The Antioch Press.

Chai, D. T. [1966] An investigation of pronominal referents in ambiguous English sentences. Report Number 14, NICH Grant HD 01368, Center for Human Growth and Development, University of Michigan.

Chomsky, N. [1957] Syntactic Structures. The Hague: Mouton.

Chomsky, N. [1963] Formal properties of grammars, in Handbook of Mathematical Psychology, Volume 2, 323-418. R. D. Luce, R. R. Bush and E. Galanter (Editors). New York: John Wiley.

Chomsky, N. [1964a] Degrees of grammaticalness, in The Structure of Language, 384-389. J. A. Fodor and J. J. Katz (Editors). Englewood Cliffs, N. J.: Prentice-Hall.

Chomsky, N. [1964b] Current issues in linguistic theory, in The Structure of Language, 50-118. J. A. Fodor and J. J. Katz (Editors). Englewood Cliffs, N. J.: Prentice-Hall.

Chomsky, N. [1965] Aspects of the Theory of Syntax. Cambridge: The MIT Press.

Chomsky, N. and Miller, G. A. [1963] Introduction to the formal analysis of natural languages, in Handbook of Mathematical Psychology, Volume 2, 269-321. R. D. Luce, R. R. Bush and E. Galanter (Editors). New York: John Wiley.

Edmundson, H. P. (Editor) [1961] Proceedings of the National Symposium on Machine Translation. Englewood Cliffs, N. J.: Prentice-Hall.

Eifermann, R. R. [1961] Negation: a linguistic variable. Acta Psychologica, 18, 258-273.

Epstein, W. [1961] The influence of syntactic structure on learning. The American Journal of Psychology, 71, 80-85.

Ervin, Susan [1964] Imitation and structural change in children's language, in New Directions in the Study of Language, 163-189. E. Lenneberg (Editor). Cambridge: The MIT Press.

Feller, W. [1957] An Introduction to Probability Theory and Its Applications, Volume 1, second edition. New York: John Wiley.

Fodor, J. A. and Katz, J. J. (Editors) [1964] The Structure of Language: Readings in the Philosophy of Language. Englewood Cliffs, N.J.: Prentice-Hall.

Fraser, C., Bellugi, Ursula, and Brown, R. [1963] Control of grammar in imitation, comprehension, and production. Journal of Verbal Learning and Verbal Behavior, 2, 121-135.

Garner, W. R. and McGill, W. J. [1956] The relation between information and variance analysis. Psychometrika, 21, 219-228.

Garvin, P. L. (Editor) [1963] Natural Language and the Computer. New York: McGraw-Hill.

Glaze, J. A. [1928] The association value of nonsense syllables. Journal of Genetic Psychology, 35, 255-269.

Gleason, H. A., Jr. [1964] The organization of language: A stratificational view, in Monograph Series on Language and Linguistics, Number 17, 75-95. C.I.J.M. Stuart (Editor). Washington, D. C.: Georgetown University Press.

Harvard University Press [1955] Tables of the Cumulative Binomial Probability Distribution. Cambridge: Harvard University Press.

Hays, D. G. [1963] Research procedures in machine translation, in Natural Language and the Computer, 183-214. P. Garvin (Editor). New York: McGraw-Hill.

Hill, A. A. [1960] Grammaticality. Word, 17, 1-10.

Hull, R. R. and Gough, P. B. [1966] Semantic versus syntactic classification of strings. Paper read at the Meetings of the Midwestern Psychological Association in Chicago, May, 1966. (Mimeo, Indiana University).

Kaplan, A. [1949] An experimental study of ambiguity and context. The RAND Corporation. P - 187. Published in Mechanical Translation, 2, 39-46, [1955].

Katz, J. J. and Fodor, J. A. [1963] The structure of a semantic theory. Language, 39, 170-210.

Katz, J. J. and Postal, P. M. [1964] An Integrated Theory of Linguistic Descriptions. Cambridge: The MIT Press.

Klima, E. S. [1964] Negation in English, in The Structure of Language, 246-323. J. Fodor and J. Katz (Editors). Englewood Cliffs, N. J.: Prentice-Hall.

Krueger, W. C. F. [1934] The relative difficulty of nonsense syllables. Journal of Experimental Psychology, 17, 145-153.

Kuno, S., Oettinger, A. G. [1963] Syntactic structure and ambiguity of English. Proceedings Fall Joint Computer Conference, 24, 397-418. Baltimore: Spartan Books.

Lamb, S. M. [1964a] The sememic approach to structural semantics, in Transcultural Studies in Cognition, American Anthropologist Supplement, 66, 57-78. A. K. Romney and R. G. D'Andrade (Editors).

Lamb, S. M. [1964b] On alternation, transformation, realization and stratification, in Monograph Series on Language and Linguistics, Number 17, 105-122. C.I.J.M. Stuart, (Editor). Washington, D.C.: Georgetown University Press.

Lamb, S. M. [1965a] The use of semantic information for the resolution of syntactic ambiguity, in Actes du Colloque International de Linguistique Appliquee. B. Pottier (Editor). Faculte des Lettres et des Sciences Humaines, Nancy, France.

Lamb, S. M. [1965b] The nature of the machine translation problem. Journal of Verbal Learning and Verbal Behavior, 4, 196-210.

Lees, R. B. [1965a] Two views of linguistic research. Linguistics, 11, 21-29.

Lees, R. B. [1965b] On the testability of linguistic predicates. Linguistics, 12, 37-48.

Lees, R. B. and Klima, E. S. [1963] Rules for English pronominalization. Language, 39, 17-28.

McGill, W. J. [1954] Multivariate information transmission, Psychometrika, 19, 97-116.

McNeill, D. [1965] Developmental Psycholinguistics. Preprint of a chapter in Handbook of Social Psychology. (Mimeo, Harvard University, Center for Cognitive Studies).

Mehler, J. [1963] Some effects of grammatical transformations on the recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 2, 346-351.

Mehler, J. and Miller, G. A. [1964] Retroactive interference in the recall of simple sentences. British Journal of Psychology, 55, 295-301.

Miller, G. A. [1956] The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 63, 81-97.

Miller, G. A. [1962] Some psychological studies of grammar. American Psychologist, 17, 748-762.

Miller, G. A. [1965] Some preliminaries to psycholinguistics. American Psychologist, 20, 15-20.

Miller, G. A. and Chomsky, N. [1963] Finitary models of language users, in Handbook of Mathematical Psychology, Volume 2, 419-491. R. D. Luce, R. R. Bush and E. Galanter (Editors). New York: John Wiley.

Miller, G. A. and Isard, S. [1964] Free recall of self-embedded English sentences. Information and Control, 7, 292-303.

Miller, G. A. and Madow, W. G. [1954] On the maximum likelihood estimate of the Shannon-Wiener measure of information, Report AFCRC-TR-54-75. Operational Applications Laboratory, ARDC.

Mood, A. M. [1950] Introduction to the Theory of Statistics. New York: McGraw-Hill.

Posner, M. I. [1963] Immediate memory in sequential tasks. Psychological Bulletin, 60, 333-349.

Posner, M. I. [1964] Information reduction in the analysis of sequential tasks. Psychological Review, 71, 491-504.

Postal, P. M. [1964] Constituent structure: A study of contemporary models of syntactic description. International Journal of American Linguistics, 30, Part III.

Quine, W. V. [1964] The problem of meaning in linguistics, in The Structure of Language, 21-32. J. A. Fodor and J. J. Katz (Editors). Englewood Cliffs, N. J.: Prentice-Hall.

Roberts, K. [1966] The interaction of normative associations and grammatical factors in sentence retention. Paper read at the Meetings of the Midwestern Psychological Association in Chicago, May, 1966. (Ditto, University of Michigan, Human Performance Center).

Robinson, Jane, J. [1962] Preliminary codes and rules for the automatic parsing of English. Memorandum RM-3339-PR, The RAND Corp.

Rosenberg, S. [1966] Recall of sentences as a function of syntactic and associative habits. Journal of Verbal Learning and Verbal Behavior, 5, 392-396.

Savin, H. B. and Perchonock, Ellen [1965] Grammatical structure and the immediate recall of English sentences. Journal of Verbal Learning and Verbal Behavior, 4, 348-353.

Shannon, C. and Weaver, W. [1949] The Mathematical Theory of Communication. Urbana, Ill.: University of Illinois Press.

Simmons, R. F. [1965] Answering English questions by computer: a survey. Communications of the Association for Computing Machinery, 8, 53-70.

Slobin, D. [1966] Grammatical transformation in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior, 5, 219-227.

Wason, P. C. [1961] Response to affirmative and negative binary statements. British Journal of Psychology, 52, 133-142.

Wilks, S. S. [1962] Mathematical Statistics. New York: John Wiley.

Yates, F. [1934] Contingency tables involving small numbers and the χ^2 test. Supplement to the Journal of the Royal Statistical Society, 1, 217-235.

Zipf, P. [1964] On understanding "understanding utterances", in The Structure of Language, 390-399. J. A. Fodor and J. J. Katz (Editors). Englewood Cliffs, N. J.: Prentice-Hall.

APPENDIX A

Test Sentences

February 8, 1966

Name: ______________________

Age: ____years______months

Major area of study: ________

INSTRUCTION

In the following pages, you will find sentences which use the pronoun he to refer to either of two persons, John or Bill. Your task is to decide, for each sentence, the most appropriate referent or antecedent of the pronoun he. Please mark an "x" on top of John or Bill as the referent of he.

Please answer every sentence in the given order. Please write down the time in the given spaces.

[Instruction sheet for the college group]

[This is the first sheet of the testing booklet]

The following 2 pages contain sentences that have nonsense words in them. For example,

John bejed Bill and he also gowed him.

Your task is to decide whether the pronoun <u>he</u> refers to <u>John</u> or to <u>Bill</u>. Please do not substitute any real words for the nonsense words, but base your judgment upon the structures of the 2 clauses. Please then mark an "x" on top of that word.

Start time: ______________

Finish time: ______________

[Instruction sheet for the college group]

[This sheet precedes the part with nonsense verb-pairs]

The following 5 pages contain sentences that you can understand easily. Please read each sentence and decide whether the pronoun he refers to John or to Bill. Then mark an "x" on top of that word.

Start time: ____________________

Finish time: ____________________

[Instruction sheet for the college group]

[This sheet precedes the part with real verb-pairs]

Name: ____________________

Birthday: ____________________

INSTRUCTIONS

In the following pages, you will find sentences using the word he in them. For example:

"John saw Bill and he called him."

Some of you may think the word he means John and some others may think the word he means Bill. Your job is to decide, for each sentence, which person (John or Bill), goes with the word he. Please mark an "x" on top of John or Bill, depending on which one you think it is.

Please answer every sentence in the given order.

[Instruction sheet for the children groups]

[This is the first sheet of the testing booklet]

The following 5 pages contain sentences that you can understand. Please read each sentence and then decide on whether the word <u>he</u> means <u>John</u> or <u>Bill</u>. Then mark an "x" on top of that word.

Start time: ________________

Finish time: ________________

[Instruction sheet for the children groups]

[This sheet precedes the part with real verb-pairs]

The sentences on the following 2 pages have made-up words in them. For example,

"John bejed Bill and he also gowed him."

I don't want you to guess some real words for the made-up words. Instead, decide which name, John or Bill, goes with the word he by understanding the sentence with the made-up words in it. Then mark an "x" on top of that name.

Start time: ________________

Finish time: ________________

[Instruction sheet for the children groups]

[This sheet precedes the part with nonsense verb-pairs]

Bill did not tef John but he was cujed back by him.

Bill and John did not qog each other and he was not tived by him.

John wuved Bill and he was keced back by him.

Bill did not zab John but he jebed him.

John kuwed Bill and he yimed him back.

Bill and John were zoced by each other but he did not peh him.

Bill and John did not kez each other but he was qemed by him.

Bill yejed John but he did not koj him back.

John and Bill were vuced by each other and he was yuged by him.

Bill and John were not kuwed by each other and he did not luj him.

Bill and John daqed each other but he was not tijed by him.

Bill was zoced by John but he did not kej him back.

John nijed Bill but he was not caqed back by him.

Bill and John kajed each other but he did not peh him.

Bill zoved John but he did not faj him.

Bill was not vebed by John but he fubed him back.

John cojed Bill but he was not wuhed by him.

John kexed Bill and he was yaged by him.

Bill did not buv John but he was ruqed by him.

John was vided by Bill and he was gojed by him.

John and Bill did not foh each other and he did not kuy him.

Bill was not yuged by John and he did not mib him back.

Bill was not ciwed by John but he was pehed back by him.

Bill was yuked by John but he was not miped back by him.

John was kaqed by Bill and he wuhed him.

Bill did not guw John and he was not quved by him.

John did not foh Bill but he vajed him back.

Bill did not jex John and he did not weq him.

John was qaped by Bill and he was tejed back by him.

Bill and John caqed each other and he was jived by him.

Bill was not zeced by John and he did not nij him.

Bill was not guked by John and he was not vused by him.

John and Bill were not fubed by each other but he was qosed by him.

John and Bill were not jebed by each other but he sebed him.

John and Bill were not vided by each other and he was not qoned by him.

John and Bill did not gax each other but he lijed him.

John was zeved by Bill and he pehed him back.

Bill qeged John and he yeced him.

John and Bill qaped each other and he tijed him.

John was not zused by Bill but he was qiged by him.

John was not yozed by Bill but he lijed him.

John was not guhed by Bill and he was not vused back by him.

Bill and John were zuted by each other but he was not kiyed by him.

John and Bill were mibed by each other and he ruqed him.

John did not quk Bill and he was not bejed back by him.

Bill did not wib John and he did not jic him back.

Bill was yoged by John but he was not kuyed by him.

Bill was quped by John but he did not bij him.

John was teased by Bill and he was teased by him.

Bill and John did not recognize each other and he did not invite him.

Bill understood John and he answered him back.

John did not frighten Bill and he did not scare him back.

Bill did not push John but he shoved him back.

Bill was remembered by John but he was not phoned by him.

Bill did not understand John and he was not answered back by him.

Bill did not punch John and he did not punch him back.

John pushed Bill but he was not shoved back by him.

John did not recognize Bill and he did not invite him.

John was frightened by Bill and he scared him.

John was not helped by Bill but he assisted him.

Bill was not pushed by John but he was shoved by him.

Bill and John heard each other but he was not called by him.

Bill and John were not heard by each other but he was called by him.

John and Bill did not hit each other and he was not kicked by him.

John did not help Bill but he was assisted by him.

John did not hear Bill but he called him.

John was not remembered by Bill but he phoned him back.

John did not hurt Bill but he hurt him.

Bill did not help John but he was assisted back by him.

Bill was understood by John and he was answered by him.

Bill was heard by John but he did not call him back.

Bill hit John and he kicked him.

John was understood by Bill and he was answered back by him.

John and Bill were frightened by each other and he scared him.

Bill did not tease John and he was not teased by him.

John and Bill understood each other and he answered him.

Bill was not struck by John but he struck him.

Bill remembered John but he did not phone him back.

John hurt Bill but he was not hurt by him.

Bill pushed John but he was not shoved by him.

Bill struck John but he did not strike him back.

Bill was not remembered by John but he phoned him.

John was heard by Bill but he did not call him.

John was hurt by Bill but he did not hurt him.

John was not hit by Bill and he did not kick him.

Bill was not recognized by John and he was not invited back by him.

Bill and John were not pushed by each other but he was shoved by him.

John was recognized by Bill and he invited him.

John was punched by Bill and he punched him.

Bill was not frightened by John and he was not scared back by him.

Bill heard John but he was not called by him.

John did not frighten Bill and he did not scare him.

John did not hear Bill but he called him back.

Bill and John did not help each other but he was assisted by him.

John and Bill did not remember each other but he was phoned by him.

Bill frightened John and he was scared back by him.

Bill was struck by John but he was not struck by him.

John and Bill were heard by each other but he did not call him.

John was not hurt by Bill but he was hurt by him.

John was frightened by Bill and he scared him back.

Bill was helped by John but he was not assisted by him.

Bill was not punched by John and he was not punched by him.

John was teased by Bill and he was teased back by him.

John and Bill were not helped by each other but he assisted him.

Bill and John were not remembered by each other but he phoned him.

Bill did not punch John and he did not punch him.

John did not recognize Bill and he did not invite him back.

Bill and John were helped by each other but he was not assisted by him.

John and Bill were hit by each other and he was kicked by him.

John and Bill did not hear each other but he called him.

Bill did not strike John but he was struck by him.

Bill was not frightened by John and he was not scared by him.

John heard Bill but he was not called back by him.

John was helped by Bill but he was not assisted back by him.

Bill and John pushed each other but he was not shoved by him.

John and Bill were pushed by each other but he did not shove him.

John hit Bill and he kicked him back.

John recognized Bill and he was invited by him.

John was hit by Bill and he was kicked by him.

Bill understood John and he answered him.

Bill was hurt by John but he did not hurt him back.

Bill was not heard by John but he was called by him.

John frightened Bill and he was scared by him.

Bill hurt John but he was not hurt back by him.

Bill was pushed by John but he did not shove him back.

John was remembered by Bill but he was not phoned back by him.

John and Bill hit each other and he kicked him.

Bill was not helped by John but he assisted him back.

Bill did not tease John and he was not teased back by him.

Bill did not hit John and he was not kicked by him.

Bill remembered John but he did not phone him.

Bill and John frightened each other and he was scared by him.

John and Bill recognized each other and he was invited by him.

John did not remember Bill but he was phoned by him.

John did not hurt Bill but he hurt him back.

John did not remember Bill but he was phoned back by him.

Bill teased John and he teased him back.

Bill was punched by John and he punched him back.

John was pushed by Bill but he did not shove him.

Bill was not punched by John and he was not punched back by him.

John and Bill were not frightened by each other and he was not scared by him.

John and Bill were not recognized by each other and he was not invited by him.

Bill did not push John but he shoved him.

John was not understood by Bill and he did not answer him back.

John helped Bill but he did not assist him.

John punched Bill and he was punched back by him.

Bill was struck by John but he was not struck back by him.

Bill was not recognized by John and he was not invited by him.

Bill did not understand John and he was not answered by him.

John and Bill were understood by each other and he was answered by him.

John was not teased by Bill and he did not tease him back.

Bill did not strike John but he was struck back by him.

Bill was not heard by John but he was called back by him.

John was not hurt by Bill but he was hurt back by him.

Bill and John did not frighten each other and he did not scare him.

Bill recognized John and he was invited back by him.

Bill helped John but he did not assist him back.

Bill and John did not understand each other and he was not answered by him.

John and Bill helped each other but he did not assist him.

John and Bill were recognized by each other and he invited him.

John and Bill remembered each other but he did not phone him.

Bill teased John and he teased him.

John and Bill were not understood by each other and he did not answer him.

Bill struck John but he did not strike him.

John was hit by Bill and he was kicked back by him.

John was not struck by Bill but he struck him back.

Bill was not hit by John and he did not kick him back.

John was not understood by Bill and he did not answer him.

John was not teased by Bill and he did not tease him.

John and Bill were remembered by each other but he was not phoned by him.

John and Bill did not push each other but he shoved him.

Bill punched John and he was punched by him.

Bill was recognized by John and he invited him back.

Bill was not pushed by John but he was shoved back by him.

Bill and John were not hit by each other and he did not kick him.

John did not hit Bill and he was not kicked back by him.

APPENDIX B

Response Matrices and Analysis by Information Transmission

Model for the Ninth Grade Ss.

Table B-1

Response matrices for 14 male[a] and 11 female ninth grade subjects

Sentence type: John and Bill Yed each other and he Xed him.

		Nonsense Verbs				Similar Verbs				Logical Reversal Verbs			
		A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
Male	A	7	4	4	7	7	7	5	7	8	7	6	5
	$\bar{A}$	7	3	5	5	4	5	7	8	3	6	4	5
	$\bar{P}$	3	4	5	5	6	5	7	6	4	6	7	7
	P	4	4	6	4	6	6	5	7	8	5	6	8
		$\hat{p}$ = .37; [1,10]				$\hat{p}$ = .44; [2,11]				$\hat{p}$ = .42; [2,11]			
Female	A	8	4	6	5	7	5	4	3	6	5	5	4
	$\bar{A}$	5	6	6	4	7	5	5	5	5	2	5	8
	$\bar{P}$	3	3	2	5	3	3	4	6	5	4	5	4
	P	5	7	3	9	6	3	5	5	4	5	4	6
		$\hat{p}$ = .46; [1,10]				$\hat{p}$ = .43; [1,10]				$\hat{p}$ = .44; [1,10]			
Total	A	15	8	10	12	14	12	9	10	14	12	11	9
	$\bar{A}$	12	9	11	9	11	10	12	13	8	8	9	13
	$\bar{P}$	6	7	7	10	9	8	11	12	9	10	12	11
	P	9	11	9	13	12	9	10	12	12	10	10	14
		$\hat{p}$ = .41; [4,16]				$\hat{p}$ = .44; [5,18]				$\hat{p}$ = .43; [5,18]			

a: Only 13 male Ss for the nonsense verbs

Table B-2

Response matrices for 14 male[a] and 11 female ninth grade subjects

Sentence type: John Yed Bill and he Xed him back.

		Nonsense Verbs				Similar Verbs				Logical Reversal Verbs				Identical Verbs			
		A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P	A	$\bar{A}$	$\bar{P}$	P
Male	A	9	11*	4	7	8	8	6	7	6	6	7	9	10	10	7	8
	$\bar{A}$	4	7	4	5	6	7	7	5	7	5	9	6	14*	9	7	3*
	$\bar{P}$	5	6	5	5	7	6	8	5	6	6	2*	4	6	6	8	8
	P	4	7	3	7	8	6	8	8	10	10	7	7	6	9	9	6
		$\hat{p}$ = .45; [2,11]				$\hat{p}$ = .49; [3,12]				$\hat{p}$ = .48; [3,12]				$\hat{p}$ = .56; [4,13]			
Female	A	6	5	6	5	9*	4	2	2	3	0*	5	5	10*	9	1*	1*
	$\bar{A}$	5	7	2	1*	8	6	4	5	4	5	5	4	6	8	2	4
	$\bar{P}$	3	1*	5	2	3	3	4	4	6	5	2	2	3	3	6	5
	P	3	3	4	4	3	4	5	4	6	6	3	2	5	2	4	8
		$\hat{p}$ = .35; [1,9]				$\hat{p}$ = .40; [1,9]				$\hat{p}$ = .36; [1,9]				$\hat{p}$ = .44; [1,10]			
Total	A	15	16*	10	12	17	12	8	9	9	6	12	14	20*	19*	8	9
	$\bar{A}$	9	14	6	6	14	13	11	10	11	10	14	10	20*	17	9	7*
	$\bar{P}$	8	7	10	7	10	9	12	9	12	11	4*	6	9	9	14	13
	P	7	10	7	11	11	10	13	12	16	16	10	9	11	11	13	14
		$\hat{p}$ = .40; [4,16]				$\hat{p}$ = .45; [6,18]				$\hat{p}$ = .44; [5,18]				$\hat{p}$ = .51; [7,19]			

a: Only 13 male Ss for the nonsense verbs
Starred cell (*) has $p < .01$

Table B-3

Response matrices for 14 male[a] and 11 female ninth grade subjects

Sentence type : John Yed Bill and he Xed him.

Male

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	5	5	6	4
$\bar{A}$	7	7	6	5
$\bar{P}$	3	4	9	3
P	8	5	7	8

$\hat{p} = .44$; [2,11]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	7	10	3*	10
$\bar{A}$	9	9	7	4
$\bar{P}$	8	4	9	7
P	7	6	8	12

$\hat{p} = .54$; [3,13]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	2*	6	5	7
$\bar{A}$	9	2*	5	7
$\bar{P}$	6	8	6	8
P	10	8	6	7

$\hat{p} = .46$; [2,12]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	12	11	8	4*
$\bar{A}$	13*	9	6	4*
$\bar{P}$	10	4*	9	8
P	7	5	10	11

$\hat{p} = .58$; [4,13]

Female

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	6	4	5	4
$\bar{A}$	3	3	3	4
$\bar{P}$	6	4	3	3
P	4	5	5	4

$\hat{p} = .38$; [1,9]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	7	5	5	9*
$\bar{A}$	5	3	1*	5
$\bar{P}$	3	5	7	4
P	5	4	0*	6

$\hat{p} = .42$; [1,9]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	2	3	2	6
$\bar{A}$	5	3	5	0*
$\bar{P}$	2	5	2	4
P	7	5	0*	3

$\hat{p} = .31$; [0,8]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	7	5	3	3
$\bar{A}$	6	7	3	3
$\bar{P}$	1*	3	6	3
P	5	4	6	5

$\hat{p} = .40$; [1,9]

Total

Nonsense Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	11	9	11	8
$\bar{A}$	10	10	9	9
$\bar{P}$	9	8	12	6
P	12	10	12	12

$\hat{p} = .42$; [5,17]

Similar Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	14	15	8	19*
$\bar{A}$	14	12	8	9
$\bar{P}$	11	9	16	11
P	12	10	8	18

$\hat{p} = .49$; [7,19]

Logical Reversal Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	4*	9	7	13
$\bar{A}$	14	5	10	7
$\bar{P}$	8	13	8	12
P	17*	13	6	10

$\hat{p} = .39$; [4,17]

Identical Verbs

	A	$\bar{A}$	$\bar{P}$	P
A	19*	16	11	7*
$\bar{A}$	19*	16	9	7*
$\bar{P}$	11	7*	15	11
P	12	9	16	16

$\hat{p} = .50$; [7,19]

a: Only 13 male Ss for the nonsense verbs
Starred cell (*) has $p < .01$

Table B-4

Components of transmitted information for 14 male[a] and 11 female ninth grade subjects. The average I.Q. scores for the male and female Ss are 98 and 103 respectively.

		H(Z)	T(ABC)	T(AB)	T(AB)/M	T(AB)/F	T(A)	T(B)	T(C)	A(ABZ)	A(ACZ)	A(BCZ)	A(ABCZ)	χ^2_{sex}
	DF		31	15	15	15	3	3	1					16
He Xed him	Nonsense	.98	4.53	1.54	5.56	2.56	.54	.44	.34	.57	.67	.30	1.68	7.87
	Similar	1.00	11.67√	5.63*	8.57	13.46*	.64	1.14	.95	3.85	1.04	.23	3.81	17.31
	Log. Rev.	.97	11.86√	6.43*	7.41	13.59*	.77	.72	1.73*	4.94	.28	.35	3.07	12.91
	Identical	1.00	13.20√	7.80√	13.20√	7.51	.40	1.50	2.50√	5.91	.23	.62	2.04	17.91
He Xed him back	Nonsense	.97	8.95	5.16	7.68	8.95	2.28*	.90	.69	1.97	.16	1.18	1.77	8.73
	Similar	.99	5.51	2.34	1.66	8.97	.27	.56	.63	1.52	.73	.32	1.49	7.61
	Log. Rev.	.98	8.91	5.56	6.67	9.33	1.29	.37	1.07	3.90	.72	.09	1.47	7.59
	Identical	1.00	15.61√	8.55√	10.16*	20.00√	.50	1.59	1.11	6.47	.03	.52	5.40	21.07
Each other	Nonsense	.98	6.34	2.87	3.09	8.84	1.09	.44	.61	1.34	.49	.29	2.07	9.20
	Similar	.99	2.82	1.25	1.70	4.23	.16	.31	.00	.79	.19	.36	1.02	3.17
	Log. Rev.	.99	3.65	1.71	3.42	3.90	.33	.19	.01	1.19	.38	.16	1.38	4.90

a: Only 13 male Ss for the nonsense frames
Checked item (√) has $p < .001$
Starred item (*) has $p < .01$

APPENDIX C

Chart Showing the Transmitted Information T(AB)

in an Ideal Group of n Ss.

Table C-1

Chart showing the transmitted information T(AB) in percent of the output uncertainty H(Z) of one bit. The response is assumed "diagonal" with m Ss out of an ideal group of n Ss.

n \ m	1	2	3	4	5	6	7	8	9	10	11	12
54	86.7✓	77.1✓	69.0✓	61.9✓	55.5✓	49.7✓	44.4✓	39.5✓	35.0✓	30.9✓	27.0✓	23.6✓
41	83.5✓	67.0✓	62.2✓	53.9✓	44.7✓	39.9✓	34.1✓	28.8✓	24.1✓	19.9✓	16.1✓	12.8✓
28	77.8✓	62.9✓	50.9✓	40.8✓	32.3✓	25.0✓	18.9✓	13.7✓	9.4✓	6.0*	3.3	
26	76.5✓	60.9✓	48.4✓	38.1✓	29.4✓	22.1✓	16.0✓	11.0✓	7.0✓	3.9		
21	72.4✓	54.6✓	40.8✓	29.8✓	20.8✓	13.7✓	8.2✓	4.1				
20	71.4✓	53.1✓	39.0✓	27.8✓	18.9✓	11.9✓	6.6					
14	62.9✓	40.8✓	25.0✓	13.7✓	6.0							
13	60.9✓	38.1✓	22.1✓	11.0*	3.9							
9	49.7✓	23.6✓	8.2									

n \ m	13	14	15	16	17	18	19	20	21	22
54	20.4✓	17.4✓	14.8✓	12.3✓	10.1✓	8.2✓	6.4✓	4.9✓	3.6✓	2.5
41	9.9✓	7.4✓	5.4✓	3.5*	2.1					

Checked cell (✓) has $p < .001$
Starred cell (*) has $p < .01$